08-09-04

AF 1120

Docket No. 46745 (48340)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

APPLICANT:	J. Weidanz, et al.		
SERIAL NO.:	08/813,781	GROUP:	1644
FILED:	March 7, 1997	EXAMINER:	R. Schwadron

FOR: FUSION PROTEINS COMPRISING BACTERIOPHAGE COAT PROTEIN AND A SINGLE-CHAIN T CELL RECEPTOR

Mail Stop Appeal Brief – Patents
Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

CERTIFICATE OF EXPRESS MAIL

I, Patricia A. Barnes, hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being deposited with the United States Postal Service with sufficient postage as "Express Mail Post Office To Addressee" service, **Label No. EV438994656US** in an envelope addressed to: Mail Stop Appeal Brief – Patents, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450 on **August 5, 2004**.

Patricia A. Barnes

Patricia A. Barnes

Sir:

RESPONSE TO NOTIFICATION OF NON-COMPLIANCE UNDER 37 CFR 1.192(c)

In response to the Notification Of Non-Compliance With 37 CFR 1.192(c) dated May 10, 2004 ("Notification"), Appellants hereby submit a revised Brief on Appeal.

As an initial matter, the undersigned expresses thanks to Dr. Schwadron for discussing the Notification in a telephone conference (July 29, 2004) including its statement that **Exhibit A** of the Appeal Brief submitted August 23, 2003 contained "new evidence" as defined in MPEP 1207.

Appellant respectfully disagrees that **Exhibit A** is "new evidence" under the cited MPEP section. For instance, MPEP1207 is entitled "Amendment Filed With or After Appeal". **Exhibit A** is not an amendment. Instead, it is merely a drawing that shows for the sake of convenience to the Board, a TCR heterodimer and single-chain TCR. It is not seen how MPEP 1207 (relating to amendments) supports the Office's position that a drawing provided for convenience is "new evidence".

However, to speed consideration of the Appeal Brief on the merits, Appellant agreed during the conference to withdraw **Exhibit A**. The enclosed Appeal Brief has been modified along this line.

Appellant respectfully points out that **Exhibit A** has been part of Appellants' appeal for more than a year. During this time, Appellants received two notifications of alleged non-compliance and in no case did those communications refer to **Exhibit A**. Appellants would have preferred to address any "new evidence" issues earlier with the USPTO. However, the Examiner did mention during the conference that if Exhibit A is removed, the Appeal Brief would be in full compliance with the requirements of 37 CFR 1.192(c). Appellants thank the Examiner for this consideration.

To avoid all doubt, the undersigned contacted SPE Christina Chan to discuss the Notification. Appellant was assured that all outstanding compliance issues would be raised during the telephone conference with Dr. Schwadron.

Accordingly, it is believed that the revised Brief (now without Exhibit A) is in full compliance with the rules. Consideration on the merits is now requested. However, if

the Office believes the Appeal Brief is still non-compliant, the undersigned would most appreciate a telephone call to discuss further issues.

Included with the present submission is a Petition for Extension of Time for two months, ie. from June 10, 2004 to August 10, 2004. Also enclosed is the requisite petition fee

Although it is not believed that any fee is needed to consider this submission, the Commissioner is hereby authorized to charge any fees which may be required to Deposit Account No. **04-1105**.

Respectfully submitted,

Date: 5 May 2004

By: _____
Robert L. Buchanan (Reg. No. 40,927)
EDWARDS & ANGELL, LLP
P. O. Box 55874
Boston, MA 02205
Tel. (617) 439-4444
Fax (617) 439-4170 / 7748
Customer No.: 21874

Doc. 453674

Docket No. 46745 (48340)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

APPLICANT: J. Weidanz, et al.

SERIAL NO.: 08/813,781 GROUP: 1644

FILED: March 7, 1997 EXAMINER: R. Schwadron

FOR: FUSION PROTEINS COMPRISING BACTERIOPHAGE COAT PROTEIN AND A SINGLE-CHAIN T CELL RECEPTOR

THE HONORABLE COMMISSIONER OF PATENTS AND TRADEMARKS
WASHINGTON, D.C. 20231

APPELLANTS' BRIEF ON APPEAL
SUBMITTED PURSUANT TO 37 C.F.R. §1.192

In support of Appellants' appeal on October 17, 2002 of the Examiner's final rejection, mailed on June 17, 2002, submitted herewith is Appellants' Brief on Appeal.

Table of Contents

REAL PARTY IN INTEREST

The real party in interest is Altor BioScience Corporation of Miramar, Florida. An assignment from the inventors to Dade International was recorded on August 18, 1997 at Reel/ Frame 8681/0081. An assignment from Dade International to Sunol Molecular Corporation was recorded on February 26, 2003 at Reel/Frame 013787/0276. An assignment from Sunol Molecular Corporation to Altor BioScience Corporation was mailed to the USPTO on February 3, 2004.

RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences known to Appellants or Appellants' representatives that will directly affect or be directly affected by or have a bearing on the Board's decision in the pending Appeal.

STATUS OF THE CLAIMS

Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 stand finally rejected under 35 U.S.C. §103 and are pending on appeal. Claims 5, 6, 10-12, 16-20, 61, 65, 70, 73 and 74 have been cancelled. Claims 3, 9, 13, 15, 21-60, 62-64, 66 and 68 were withdrawn from consideration by the Examiner.

STATUS OF THE AMENDMENTS

Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 on appeal are set forth in Exhibit A hereto. There has been no further amendment to the claims.

SUMMARY OF THE CLAIMED INVENTION

The claimed invention features a soluble fusion protein engineered to include a bacteriophage coat protein fused to a single-chain T cell receptor ("scTCR"). The single-chain T cell receptor was itself designed to include an alpha-variable region ("V-α") fused to a beta-variable region ("V-β"). The single-chain T cell receptor forms a pocket that binds antigen when the antigen. The claimed soluble fusion protein further includes a beta-constant region ("C-β") region that can be fused to V-β, for example.

T cells help defend the body against infection. The cells have membrane bound receptors that bind foreign antigen with the assistance of a protein complex called "MHC". A key receptor is called the T cell receptor ("TCR"). The chemical structure and function of the TCR has been extensively studied. For instance, it is known that formation of a TCR-antigen-MHC complex is an important step toward fighting infection.

Appellants discovered that by adding a bacteriophage coat protein to the scTCR, it is possible to produce a fully soluble and functional scTCR. Unlike prior scTCRs, the claimed fusion proteins were found to be fully soluble, functional, and obtainable in significant quantities without difficulty. The claimed fusion proteins have a wide spectrum of important uses as described throughout the instant patent application.

See the Summary Of The Invention at pg. 3, line 22 to pg. 11, line 21. See also the Background at pg. 1, line 12 to pg. 3, line 19 for related information. None of the art of record in this case shows an attempt to make a scTCR that includes a fused bacteriophage coat protein.

ISSUE

This appeal presents the issue of whether the Examiner erred in rejecting claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 under 35 U.S.C. § 103 in view of Chung, S. et al. (1994) *Proc. Natl. Acad. Sci. (USA)* 91: 12654 in view of U.S Pat. No. 5,759,817 to Barbas, Onda, T et al. (1996) *Mol. Immunol.* 32: 1387; and Huse et al. (1992) *J. Immunol.* 149: 3914. Appellants will refer to these citations as "Chung", "Barbas", "Onda" and "Huse", respectively; unless stated otherwise. There are no other pending rejections of record in this case.

GROUPING OF THE CLAIMS

All of claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 stand or fall together for the purpose of the present appeal.

CASE HISTORY

In consideration of the Examiner's position in this case, Appellants have summarized the prosecution history with respect to the 35 U.S.C. §103 rejection at issue. A more detailed discussion of Appellants' rebuttal to those arguments will follow under Argument.

A. A patent application was filed on March 7, 1997 with 59 claims and assigned Serial No.: 08/813,781 by the USPTO.

B. An Office Action was mailed to Appellants on August 25, 2000 by Examiner Schwadron. The Examiner stated the following with respect to the §103 rejection on appeal:

> Claims 1, 2, 4, 7, 8, 14, 67, 69, 71, 72 are rejected under 35 U.S.C. 103(a) as unpatentable over Chung et al. in view of Barbas US 5,759,817 (filed Jan. 27, 1992), Onda et al. (Molecular Immunology 32:1387, 1995), and Huse et al. J. Immunology 149:3914. 1992
>
> Chung et al. teaches a single chain T cell receptor which specifically binds to peptide ligand (see abstract). Chung et al. further teaches one embodiment of human single chain TCR in which C-terminus of V α domain is linked to N-terminus of V β chain via a 15 amino acid residue flexible amino acid liner and the C-terminus of the V β chain is linked to the beta chain constant domain (see Figure 1). In one embodiment the C terminus of V β chain is linked to a alkaline phosphatase (PI) protein tag (see page 12655). Chung et also teach that the purpose of the linker is to enhance the binding characteristics of the soluble T cell receptor and that linkers of about 10 to 30 amino acid residues would be considered to be sufficient. Chung et al. teach that the TCR fusion protein can bind antigenic protein, thus teaching that the TCR fusion protein comprises an antigen binding pocket. Chung et al. teaches a TCR fusion protein comprising V-α-peptide linker-V β-C β linked to GPI anchor and expression of such a fusion protein in a transfected eukaryotic cell (see results section). Chung et al. disclose that the soluble form of TCR protein could be readily obtained by enzymatic cleavage with phosphatidylinostol-specific phospholipase C (PI-PLC) (see page 12656). Chung et al. teaches expression of said TCR fusion protein in a bacterial cell system in which the N terminus of the Cβ region is linked to a histidine protein tag. Chung et al. also disclose a scTCR in which comprises V-α-peptide linker-V β-C β GPI in which the Cβ component consists of the β chain sequence ending right before the last cysteine (the sixth cysteine) (see page 12655). Chung et al. further teach that TCR fusion proteins which do not contain the CB do not fold into the native conformation. The scTCR disclosed by Chung et al. meet the length limitations of the Vα. and Vβ region recited in claims 69 and 71. Chung et al. teach a soluble fusion protein comprising a Vα-peptide linker-Vβ-Cβ fragment-protein tag (eg. GPI). **Chung et al. does not teach a TCR fusion protein further comprising bacteriophage VIII coat protein.**

However, Barbas discloses a soluble fusion protein comprising a bacteriophage coat protein fragment covalently linked to a single-chain heterodimeric receptor (see abstract and column 15, lines 27-28, in particular). Barbas also discloses that the fusion protein may comprise domains of heterodimeric proteins derived from several ligand binding proteins, including immunoglobulins and T cell receptors (see column 17, lines 62-66 and column 19, lines, 9-28. Barbas discloses that T cell receptor comprises alpha and beta chains each having a variable(V) and constant(C) region and T cell receptor has similarities in genetic organization and function to immunoglobulins (see column 19, lines 19-22, in particular). Barbas also teaches that bacteriophage coat protein may be derived from cpIII or cpVIII (see column 31, lines 10-28, in particular). Barbas discloses that expression vectors expressing soluble fusion proteins in which the ligand binding region is fused to bacteria coat protein allows the expression of the multiple fusion proteins on the surface of phage particles IE approximately 2700 cpVIII heterodimer receptor molecules per phage particle (see column 39 line 64 through column 40, line 7, in particular). Barbas further discloses that a short length of amino acid sequence at the amino end of a protein (IE a protein tag) directs the protein to periplasmic space (see column 8, lines 49-55, in particular. One embodiment of the invention is disclosed to be a fusion protein comprising in sequence a leader sequence-peptide linker-V region amino acid residue-peptide linker-phage coat protein and that in one embodiment, the second linker can define a proteolytic cleavage site which allows the heterodimeric receptor to be cleaved from the bacteriophage coat protein to which it is attached (see column 14, lines 60-65). **Thus Barbas discloses but does not exemplify a soluble fusion protein comprising a bacteriophage coat protein covalently linked to T cell receptor domains.**[1]

Onda et al. disclose a soluble fusion protein comprising a bacteriophage coat protein covalently linked to a single-chain T cell receptor by a peptide linker sequence wherein the single TCR chain is the alpha chain and the bacteriophage coat protein is cpVIII (see abstract and Figure 1, in particular). Onda et al. also teach that TCR-bacteriophage coat protein fusion protein can be used to study specific binding interactions of the TCR chain to antigenic ligands (see paragraph bridging pages 1394-1395, in particular).

Huse et al. teach that fusion proteins comprising a single chain fusion protein comprising Fab fragment of immunoglobulin (which comprises the antigen binding pocket of the immunoglobulin molecule) and bacteriophage VIII coat protein can be produced and display the fusion protein when expressed in a M13 derived vector. Huse et al. further teach that bacteriophage VIII coat protein

[1] Examiner Schwadron withdrew this statement in the next Office Action dated June 17, 2002 (see below). The statement was originally made by a prior Examiner (Lubet) in a related §103 rejection that has been withdrawn. In that earlier rejection, Lubet argued that Barbas and Onda **do not teach** a soluble fusion protein in which a single-chain TCR linked to a bacteriophage coat protein. See the Office Action dated June 23, 1998 at pg. 8, part B. In the §103 rejection on appeal, Examiner Schwadron argued that Barbas and Onda teach use of TCR-bacteriophage VIII fusion protein. That molecule, a heterodimeric T cell receptor fusion protein, is not the claimed invention.

fusion protein can recovered from culture medium or from the periplasmic space (see abstract).

Therefore it would have been prima facie obvious to one with ordinary skill in the art at the time the invention was made to make a soluble TCR fusion protein comprising the Vα..-peptide linker-Vβ -Cb fragment-protein taught by Chung et al. linked to a bacteriophage VIII coat protein because **Barbas et al. and Onda et al. teach TCR-bacteriophage VIII coat fusion proteins can be used to study antigen binding properties of such a fusion protein** and Huse et al. teach that fusion proteins comprising bacteriophage VIII coat protein can be produced in bacteria and recovered in relatively large quantities.

One with skill in the art would be motivated to make such a fusion protein to study the antigen binding region of the TCR component or to use the protein to elicit anti-idiotypic antibodies. One with skill in the art would be motivated to make such a fusion protein in which the Vα and Vβ region was derived from human TCR in order to study human TCR properties or to elicit anti-idiotypic antibodies to the TCR component of the protein.

The preceding grounds of rejection have been maintained since the August 25, 2000 Office Action despite Appellants' rebuttal argument and claim amendments discussed below.

ℂ. On October 25, 2000, Appellants' representative met with Examiner Schwadron at the USPTO and discussed the art cited. No agreement was reached.

𝔻. Appellants submitted a response to the rejection set forth in paragraph B, above, on February 22, 2001. In that response, Appellants rebutted the <u>prima facie</u> obviousness argument by pointing out: **1)** that the cited references did not teach or suggest that the membrane "anchor" of Chung's single-chain TCR (GPI: a membrane protein) could be substituted with the bacteriophage coat protein of Barbas' TCR; **2)** that there was no reasonable expectation that the substitution (switching Chung's anchor for Barbas' phage coat protein) could be achieved in view of substantial differences between scTCRs and TCRs; and **3)** that the Examiner's citation of Onda was not correct ie., it does not disclose TCR-bacteriophage coat protein fusions, but instead, dwelt on smaller constructs having only a V-α chain (but no V-β chain). Onda characterized his constructs as having "unusual" binding properties that were not characteristic of TCRs. A subset of such constructs were reported by Onda not to work at all.

Appellants also discussed the **Holler** reference: a peer-reviewed scientific article from the U.S. Academy of Sciences (*PNAS (USA)* (2000) 97: 5387 at 5389). Holler provided independent and objective evidence of the **long-felt need and failure of others** in the field to make and use the claimed fusion molecules. Specifically, Holler stated that phage display had not yet proven successful in the engineering of scTCRs.[2]

E. In response to the Appellants' arguments in paragraph D above, the Examiner issued a Final Office Action dated June 17, 2002.[3] The Examiner maintained the prima facie rejection and stated:

> Regarding appellants comments, while **heterodimeric molecules are a preferred embodiment** disclosed in Barbas et al., Barbas et al. disclose: "In another embodiment, the present invention contemplates a polypeptide comprising an insert domain flanked by an amini-terminal secretion signal domain and a carboxy-terminal filamentous phage coat protein membrane anchor domain." (column 14, first complete paragraph).
>
> Barbas et al. further disclose than said construct could include a **"receptor protein"** (column 14, second paragraph), indicating that the disclosed method could be used for receptors per se (eg. single chain or heterodimeric or single chain heteromers). Single chain T cell receptors were known in the art (see Chung et al.).
>
> Regarding appellants comments about the single chain TCR taught by Chung et al.,
>
> Chung et al. teach that the GPI anchor is cleaved and the soluble TCR still has all the antigen binding properties of the TCR (see pages 12656-12658). Thus, the GPI anchor is not required for the soluble TCR to function, it is just used in one particular method of making the soluble TCR. Regarding motivation to create the claimed invention, **Chung et al. discloses that it would be desirable to produce their TCR in a phage display system (see page 12658, first column).** In addition, Barbas et al. teach the advantages of their system for the production

---

[2] By "phage display" is meant the process of making a recombinant bacteriophage expressing the scTCR as part of the phage protein coat. After infecting bacteria with a recombinant phage engineered to produce the scTCR, the protein would be "displayed" on the bacterial cell surface as bacteriophage. The scTCR "displayed" in this manner would be amendable to engineering.

[3] The Examiner essentially repeated his rejection of the claims as set forth in the August 25, 2000 rejection. But see footnote 1. With respect to Examiner Schwadron's discussion about Holler, Appellants submitted Weidanz et al. as part of a Rule 132 Declaration to address an obviousness rejection that has since been withdrawn. The reference provided evidence that a particular scTCR-bacteriophage coat protein vector (pKC44) was capable of forming an antigen binding site when expressed. The reference is not prior art to the present application.

of peptides. Regarding reasonable expectation of success, both Barbas et al. and Chung et al. disclose use of phage display systems to produce single chain antibodies (see column 2, third paragraph from bottom and page 12658, first column). In addition, the soluble single chain TCR molecules functions with or without the GPI linker indicating that the construct itself is functional.

Regarding appellants comments about Holler et al., said publication was published in May 2000. In the amendment filed 6/3/2000, applicant submitted a publication by Weidanz et al. (J. Imm. Methods 1998) which discloses the claimed invention. Thus, **it appears that Holler et al. simply are not familiar with the prior art. Thus, the comments of Holler et al. carry no weight because two years prior to the Holler et al. publication, Weidanz et al. had already published data regarding the production of single chain TCR using bacteriophage**. Furthermore, Holler et al. discloses a yeast system for producing a single chain TCR and it appears that the main focus of Holler et al. is to promote their system.

Regarding appellants comments about Onda et al., the instant rejection indicates that "**Onda et al. disclose a soluble fusion protein comprising a bacteriophage coat protein covalently linked to a single-chain T cell receptor** by a peptide linker sequence wherein the single TCR chain is the alpha chain and the bacteriophage coat protein is cpVIII (see abstract and Figure 1, in particular)". The art recognizes that the alpha and beta chains of the TCR generally both are involved in antigen binding. The art also recognizes that soluble TCR which bind antigen would have a variety of uses.

F. Appellants filed a Notice of Appeal on October 17, 2002.

ARGUMENTS

As an initial mater, Appellants wish to emphasize the substantial differences between TCR heterodimers ("TCRs") and single-claim T cell receptors ("sc-TCRs").

The TCR is a heterodimer with one α chain and one β chain.[4] Each of these chains passes from the exterior of the T cell, through the cell membrane, and into the cell interior (cytosol). The α and β chains each include a variable (V-α, V-β) region that cooperate to form an antigen binding pocket. The regions are "variable" because its chemical structure can be

[4] A textbook in the field describes the TCR as "a heterodimer composed of an α and a β polypeptide chain, both of which are glycosylated." See Alberts, B et al. (1989) in *Molecular Biology of the Cell*, 2nd Ed. Garland Publishing, Inc. New York at pg. 1037. By convention, a "heterodimer" such as the TCR properly has two chains (dimer) both of which are different (hetero) from the other.

changed to make a pocket that fits another antigen. Each of the V-α and V-β regions are associated with a constant (C) region.

TCR heterodimers have been extremely difficult to isolate from T cells. This problem has hindered study of the receptor. One approach to address the problem has been to make single-chain T cell receptors ("scTCRs"). These synthetic receptors include, on one chain instead of two, a fused V-α and V-β region. It has been customary to space the V-α and V-β regions from each other with a flexible linker to allow the regions to make an antigen binding pocket. Unfortunately, many scTCRs have still proven to be difficult to make and use.

Appellants point out that "TCR" is understood in the field to mean a **heterodimeric** T cell receptor. The TCR is a membrane bound (insoluble) receptor in which the α and β chains cooperate to bind antigen. Reference to a "scTCR" is understood to mean a synthetic **single-chain** molecule that includes the V-α and V-β regions bound together usually through a flexible linker. Unlike the TCR, the scTCR binds antigen with only one chain. The TCR and scTCR are structurally distinct proteins that are different molecules that bind antigen in different ways.

I. Summary of the Cited Art

A. Chung reports functional three-domain single-chain T cell receptors consisting of a human Vα and Vβ region that recognizes a particular antigen (HLA-DR2b/myelin basic protein). Chung determined that it was important to fuse a Cβ region to the Vβ region. Such a three domain construct, when linked to a synthetic cell membrane anchor (glycosyl phosphatidyl-inositol (GPI) or CD3ζ fragment), was found to be expressed and functional. Chung disclosed that the cell membrane anchor could be cleaved from the single-chain receptors to obtain soluble protein. See the Abstract.

Chung opined that his single-chain design "may allow" construction of TCR phage libraries and that such libraries "may be" tools for studying TCRs. See pg. 12658. However, there is no specific disclosure in Chung about how such libraries could be made or, if made, whether his single-chain TCRs could tolerate fusion of the bacteriophage coat protein. Chung does not report or suggest that the GPI or CD3ζ membrane anchor could be substituted with a bacteriophage coat protein. Even if there was such a teaching, there is no disclosure in the

reference about whether a recombinant bacteriophage could tolerate Chung's scTCR as part of the phage coat.

B. Barbas discloses heterodimeric receptor libraries that use phagmids. In the Abstract, phage are taught to encapsulate a genome encoding first and second polypeptides of a receptor such as an antibody; in which the first and second polypeptides are integrated into the coat matrix of the phage. Barbas generally discloses that such phage may include a polypeptide with an "insert domain" that has a receptor domain flanked by a secretion signal domain and a phage coat protein membrane anchor domain. Col. 14, lines 10-14. Heterodimeric receptors are preferred. See Col. 3, lines 1-41; Col. 14, lines 15-29; and Col. 15, lines 28-32. According to Barbas, there was some uncertainty in the field about which portions of bacteriophage coat proteins were needed for phage assembly. Col. 2, lines 19-46.

Barbas does not teach how to make or use a scTCR with or without a fused bacteriophage coat protein.

C. Onda reports use of a phage display system to explore binding interactions between the V-α region and antigen. Onda did not disclose use of the system to study TCR or single-chain TCR interactions. In the Abstract, Onda provides at pg. 1387:

> We utilized an M13 phage display system, designed for multivalent receptor display, to explore specific binding interactions between various TCR **α chains** and specific antigen in the absence of MHC.

That is, Onda fused **only** the V-α region to bacteriophage coat protein. The constructs **do not** include a V-β region and are **not scTCR fusion proteins.** Onda's fusions are much smaller TCR "half-molecules" lacking the V-β region and antigen binding pocket of Appellants' scTCR.

Onda at pg. 1395, col. 1, cautioned that his constructs were **unusual** and not typical of standard TCR interactions:

> Our results extend these findings by demonstrating that the dominant interactions of certain TCR**α chains** for peptide antigens may be sufficiently high that they can be analysed independently. However, these interactions are **quite unusual** in that they do not require the expression of the second TCR subunit or normal MHC and coreceptor interactions. **These results may raise concern that this model does not reflect typical TCR-ligand interactions.**

Significantly, only some of Onda's Vα chain fusion proteins were reported to bind antigen when fused to bacteriophage coat protein. At pg. 1395, col. 2 he states that:

> ...**only a subset of TCR Vα have capacity for direct interactions with antigen strong enough to be detectable in this system.**

Onda does not teach or suggest fusing a scTCR (V-α and V-β) to a bacteriophage fusion protein.

D. Huse described a phage vector system for screening and producing antibody F(ab) fragments.[5] Huse's system was reportedly used to produce free F(ab) and F(ab) displayed on the surface of bacteriophage. According to Huse however, not all attempts to produce F(ab) were successful. In more than a few instances, the recombinant bacteriophage made to produce the fusions apparently would not tolerate certain amounts of antibody protein. In describing attempts to display certain antibody H and L chains with his phage vector system, Huse stated on pg. 3919, col. 2 that:

> Phage titers of [phage vector] infected cultures were found to **decrease** relative to the level of F(ab)-pVIII fusion protein incorporation (cite omitted). Taken together, these results suggest that a **functionally viable phage particle may be able to tolerate a limited number of incorporated F(ab)-pVIII fusion products and that the amount of F(ab) incorporated into the phage coat may adversely affect phage titers and overall F(ab) yield.**

Huse does not disclose fusing a bacteriophage coat protein to a scTCR or TCR.

II. Summary of the Examiner's Argument

Grounds for the present rejection under 35 U.S.C §103 were formulated in the Office Action dated August 25, 2000. See paragraph B, above. The basis for the rejection has not changed substantially in the face of Appellants' arguments and claim amendments.

According to Appellants' understanding of the Examiner's alleged prima facie case, the primary references, Chung and Barbas, are alleged to teach a scTCR linked to a bacteriophage fusion protein. Onda and Barbas are relied on to teach that TCR-bacteriophage fusion proteins can be used to study antigen binding. Huse is used to teach that fusion proteins with the coat

[5] F(ab) is an abbreviation for an antigen binding fragment of an antibody (fragment antigen binding). F(ab) is a heterodimer consisting of two different chains ie., the antibody light and heavy chain. F(ab) is readily made by cleaving whole antibodies with specific proteolytic enzymes.

protein can be made in bacteria. The foundation of the Examiner's position is that because Barbas, Onda, and Huse teach some bacteriophage coat protein fusions, then it would be obvious to make fusions with Chung's scTCRs. Although facially somewhat plausible, the rejection is flawed on both scientific and legal principles as discussed herein.

III. The Examiner Erred in Rejecting Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 as Being Obvious

A. Requirements of the prima facie case and its maintenance.

The Examiner erred in maintaining the obviousness rejection in the face of claim amendments and the state of the art as submitted made in this case and its parent. The Federal Circuit has reiterated that an Examiner's prima facie case is but a procedural tool of patent examination, with the express purpose of allocating the burdens of going forward as between the Examiner and Applicant. See In re Deckler 977 F.2d at 1449, citations omitted):

> Specifically, when obviousness is at issue, the examiner has the burden of persuasion and therefore the initial burden of production. Satisfying the burden of persuasion, constitutes a so-called prima facie showing. Once that burden is met, the applicant has the burden of production to demonstrate that the examiner's preliminary determination is not correct. The examiner, and if later involved, the Board, retain the ultimate burden of persuasion on this issue.

Clearly, as demonstrated herein, adequate evidence of the unobviousness of the claimed invention was provided by Appellants to shift the burden of persuasion to the Examiner.

In view thereof, it is requested that the Board review the obviousness question based on the invention as claimed, and the cited references, including all relevant parts thereof.

B. Standard For Reviewing An Obviousness Rejection under 35 USC §103.

The Federal Circuit has reiterated the manner in which obviousness rejections are to be reviewed. Where claimed subject matter has been rejected as obvious in view of a combination of prior art references, "a proper analysis under section 103 requires, inter alia, consideration of two factors: (1) whether the prior art would have suggested to those of ordinary skill in the art that they should make the claimed composition or device, or carry out the claimed process; and (2) whether the prior art would also have revealed that in so making or carrying out, those of

ordinary skill would have a reasonable expectation of success." In re Vaeck, 947 F.2d 488, 493, 20 U.S.P.Q.2d 1438, 1442 (Fed. Cir. 1991), cited In re Dow Chemical Co., 837 F.2d 469, 473, 5 U.S.P.Q.2d 1529, 1531 (Fed. Cir. 1988). As the Federal Circuit emphasized by succinctly summarizing: "Both the suggestion and the reasonable expectation of success must be founded in the prior art, not the Applicants' disclosure." Id. See also In re Merck & Co., Inc., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

More recently, the Federal Circuit has reviewed the case law regarding 35 U.S.C. §103. See In re Sang-Su Lee 277 F.3d 1388, 61 U.S.P.Q.2d 1430 (Fed. Cir. 2002).

Should the Board adopt the Examiner's prima facie case, Appellants submit that the claimed invention would not have been obvious in view of the legal standard summarized above.

C. **No Prima Facie Case of Obviousness**

As noted above, the foundation of the Examiner' argument rests on the belief that it would be obvious to fuse Chung's scTCR to the bacteriophage coat protein of Barbas because, allegedly, Barbas and Onda teach TCR-bacteriophage coat fusion proteins and Huse discloses that fusion protein with such a coat can be made in bacteria.

For the Examiner's prima facie case to stand, it is imperative that he establish that: 1) The cited references disclose or suggest fusing a bacteriophage coat protein to Chung's scTCR; 2) there is a settled role for the bacteriophage coat protein in making fusion proteins; and that 3) one could fuse Chung's scTCR to Barbas' coat protein with a reasonable expectation of success.

The Examiner's position is not supported by any of these points. Barbas, as relied on, does not teach or suggest a scTCR or even fusion of a scTCR to a bacteriophage coat protein. Onda and Huse, when read in their entirety, exemplify uncertainty in the field about using the bacteriophage coat protein to make certain fusion proteins. Even Barbas admitted that there was some doubt about how much one could change certain bacteriophage coat proteins without hindering phage assembly. Moreover, some of Onda's and Huse' molecules did not work well. Others did not work at all. On top of that uncertainty is heaped additional doubt about whether

Chung's "anchor" fragments could be substituted with the bacteriophage coat protein of Barbas. There was also doubt about whether the bacteriophage would tolerate fusion of the scTCR to its coat.

i) *Barbas and Onda does not teach or suggest a single-chain TCR (scTCR)*

In the Office Action dated August 25, 2000, Examiner Schwadron took the position that Barbas discloses:

> soluble fusion protein comprising a bacteriophage coat protein fragment covalently linked to a **single-chain heterodimeric receptor** (see the abstract and column 15, lines 27-28, in particular). Barbas also discloses that the **fusion protein may comprise domains of heterodimeric proteins derived from several ligand binding proteins,** including immunoglobulins and T cell receptors (see column 17, lines 62-66 and column 19, lines, 9-28. Barbas discloses that T cell receptor comprises alpha and beta chains each having a variable(V) and constant(C) region and T cell receptor has similarities in genetic organization and function to immunoglobulins (see column 19, lines 19-22, in particular).
>
> * * * *
>
> **Thus Barbas discloses but does not exemplify a soluble fusion protein comprising a bacteriophage coat protein covalently linked to T cell receptor domains**

The heterodimeric receptor proteins pointed out by the Examiner are not scTCRs. Heterodimeric proteins, and particularly the TCR of Barbas, are understood in the field to consist of two different α and β chains. Brief at pg. 1. Unlike the TCR, the scTCR of Appellants' claimed invention is a single-chain molecule with a V-α chain fused to a V-β chain. The position that Barbas discloses a "single-chain heterodimeric" receptor simply makes no sense. How can a single-chain molecule be a "heterodimer" when that requires two (dimer) different (hetero) chains? Brief at pg. 1 and footnote 1. Barbas could not have had the single-chain constructs of Onda and Chung in mind. Those references were published well after the priority date of the Barbas patent. Thus, nowhere in the reference is there any disclosure about how to make or use a scTCR.

Faced with this rebuttal, the Examiner took the position in the Final Office that Barbas' disclosure of "polypeptides comprising an insert domain" and "receptor proteins" should be read to include Chung's scTCRs. Also included in that sweeping reading of Barbas are "single chain or heterodimeric or single chain heteromers". That position is without merit. Too much is read from Barbas. It does not provide for any scTCR molecules. If the Examiner's overly-broad view

of the patent is allowed to stand and sweep in scTCRs, even though Chung's were published well after Barbas' priority date, it will preempt any attempt to obtain patent protection for scTCR-bacteriophage coat fusion proteins. A principle focus of Barbas was to provide **heterodimeric** receptors linked to a phage coat protein. See the Title of the patent, the Abstract and col. 3, lines 1-41, for instance. Such receptors are not the fusion proteins Appellants claim and there is no suggestion in Barbas to make or use them.

Even assuming, *arguendo*, that the Examiner is correct and that Barbas taught or suggested a scTCR (years before Chung or Onda were published), one reading Barbas in that way would be confused in light of the accepted understanding in the field that a heterodimer such as the TCR is a complex of two different polypeptide chains. Brief at pp. 1 – 2 and footnote 1.

As captioned above, Onda does not disclose a TCR or scTCR fusion to bacteriophage coat protein as alleged by the Examiner in the August 25, 2000 and June 17, 2002 Office Actions. Instead, Onda reports fusion of **TCR α chains** to bacteriophage coat protein. The TCR α chain is merely a part of the larger scTCR Appellants worked with. That is, the prior constructs are significantly smaller (and less likely to cause solubility problems when fused to coat proteins) than the scTCR fusions Appellants successfully made.

Moreover, the Examiner ignored Onda's clear hesitation about reading too much from TCR α chain constructs that include a fused bacteriophage coat protein. According to Onda, the interactions of the constructs were **unusual and not typical of TCR-ligand interactions.** See above and Onda at pg. 1395, col. 1.

Importantly, only some of Onda's TCR α chain constructs even worked to bind antigen. See above and pg. 1395 of Onda at col. 2, second full paragraph.

According to Onda then, some TCR α chain-bacteriophage coat protein fusions work and some do not. Those that do work were viewed as "unusual" and "not typical". In view of this caution, one working in this field would not be encouraged to fuse a bacteriophage coat protein

to a scTCR. None of the other cited references shed any light on Onda's clear hesitation to extend there findings to other TCR molecules.

The Examiner thus erred in trying to formulate a prima facie case by not giving due weight to all relevant portions of Onda. Contrary to this practice, it is well established that the Examiner must consider all relevant portions of cited references, including those portions which substantially weaken her position. In particular, the former CCPA stated in In re Mercier 515 F.2d 1161, 185 USPQ at 778:

> The relevant portions of a reference include not only those teachings which would suggest particular aspects of an invention to one having ordinary skill in the art, but also those teachings which would lead such a person away from the claimed invention.

See also Phillips Petroleum Co. v. U.S. Steel Corp., 673 F.Supp. 1278, 1315, 6 USPQ2d 1065, 1093 (D.Del. 1987), *aff'd*, 865 F.2d 1247, 9 USPQ2d 1461 (Fed. Cir. 1989).

The Board is thus urged to take Onda in its entirety and to consider all relevant portions of it including the passages quoted above. Read in this way, as it should, the reference would lead one in this field to doubt whether it would be feasible to fuse a bacteriophage coat protein to a scTCR to produce a soluble and functional fusion protein.

ii) *Huse reported difficulties producing some bacteriophage coat protein fusions*

The Huse reference, as quoted above, reported that not all F(ab)-pVIII (bacteriophage) coat proteins could be made at high titre. That is, Huse stated that the bacteriophage may not tolerate some amounts of F(ab) constructs, thereby decreasing phage titres and overall F(ab) yield. See above and Huse at pg. 3919, col. 2. When Huse is read in its entirety, as it should, the Examiner's statement that "Huse et al. teach that fusion proteins comprising bacteriophage VIII coat protein can be produced in bacteria" is an unsupported generalization. Huse clearly found that some amounts of heterodimeric F(ab) constructs harmed the bacteriophage that carried them. In view of this warning, a worker in the field would have good reason to doubt whether a bacteriophage could be fused to a scTCR or even a heterodimer such as a TCR without considerable experimentation.

The Examiner took the position in the Final Office Action that Barbas and Chung provide a reasonable expectation that one could make the claimed fusion proteins:

> Regarding **reasonable expectation of success**, both Barbas et al. and Chung et al. disclose use of phage display systems to produce single chain antibodies (see column 2, third paragraph from bottom and page 12658, first column).

However as clearly illustrated by Huse, not all phage display systems using antibodies work as expected. Some amounts of heterodimeric F(ab) antibodies cause problems. Thus the Examiner's position is not supported by the art of record in this case.

The Board is thus requested to take Huse in its entirety and to consider all relevant portions of it including the passage quoted above. In re Mercier 515 F.2d 1161, 185 USPQ at 778; and Phillips Petroleum Co. v. U.S. Steel Corp., 673 F.Supp. 1278, 1315, 6 USPQ2d 1065, 1093 (D.Del. 1987), *aff'd*, 865 F.2d 1247, 9 USPQ2d 1461 (Fed. Cir. 1989).

The substantial uncertainties raised by Onda and Huse have not been addressed by the Examiner. No objective scientific work has been made of record to resolve or explain them. Read in their entirety, as they should, Onda and Huse point out problems about making and using some bacteriophage coat protein fusions. Even if one skilled in this field were to read Onda and Huse selectively and disregard their warnings, there is still nothing in the art relied on to suggest that one could make or use a scTCR fusion to a bacteriophage coat protein. Even Barbas admitted that there was uncertainty about what coat protein parts could be manipulated for phage assembly.

In marked contrast, Appellants have demonstrated that it is possible to fuse a bacteriophage coat protein to the scTCR and obtain fully soluble and functional fusion protein. See Appellants' patent specification at Example 1 (showing construction of soluble scTCR fusion proteins); Examples 2-3 (production of special vectors to make the scTCR fusion proteins); Example s 4-5 (expression of Appellants' soluble scTCR fusions); Example 6 (purification of the soluble scTCR fusion proteins); Examples 7-11 and 16 (characterization of particular scTCR fusion proteins); and Example 15 (analysis of a bacteriophage library expressing Appellants' scTCR).

Accordingly, the §103 rejection fails both prongs of the Federal Circuit test for determining obviousness. See In re Vaeck, *supra*; and In re Dow Chemical Co., *supra*. It is submitted that the Board reverse the obviousness rejection in light of this test.

Whether or not the Examiner is taking the position that it would be **obvious to try** to make the claimed scTCR fusion proteins, both the Board and Federal Circuit have made it quite clear that this is not a burden that Appellants must bear. In particular, the Court in In re O'Farrell, 7 USPQ 2d 1673 (1988) held at page 1681:

> The admonition that "obvious to try" is not the standard under §103 has been directed mainly at two kinds of error. In some cases, what would have been "obvious to try" would have been to vary all parameters or try each of numerous possible choices until one possibly arrived at a successful result, where the prior art gave either no indication of which parameters were critical or no direction as to which of many possible choices is likely to be successful.

See also Ex parte Old, 229 USPQ 196, 200 (1985).

While the Court in In re O'Farrell went on to state that while obviousness does not require absolute predictability of success, what is required under §103 is a reasonable expectation of success.

Thus whether those in the field may have been tempted to fuse a bacteriophage coat protein to a scTCR, the field would have been cautioned from doing so in view of the warnings of Onda, Huse and to some extent even Barbas. The Barbas patent does not disclose or suggest any scTCR fusion to the coat protein. Chung does nothing to remedy these defects. In short, the field's unsuccessful experience with some amounts of antibody heterodimers (Huse) and some single-chain constructs (Onda), provides at worst no basis for believing that fusion of a bacteriophage coat protein to an scTCR will work and at best, a reason to doubt that such a fusion will result in a fully soluble and functional protein.

iii) *No teaching or suggestion that Chung's "anchor" fragment could be substituted with Barbas' bacteriophage coat protein*

Chung reported TCRs linked to a cell membrane anchor (glycosysl phosphatidylinositol (GPI) or murine CD3 ζ chain). The anchor apparently helps to express the single-chain TCRs. The anchor molecules are entirely different from the coat proteins of Barbas both in terms of chemical structure and function. For example, Chung's anchors are hydrophobic cell membrane proteins while those of Barbas are relatively more hydrophilic bacteriophage coat components. Chung's anchor apparently plants the scTCR in the membrane while the coat envelops the phage.

The Examiner has pointed to no teaching or suggestion in the cited art that Chung's anchor molecules could be substituted with Barbas' bacteriophage coat proteins. The obviousness rejection falls far short of establishing any nexus between Chung's anchors, which are attached to his scTCRs, and the coat proteins reported by Barbas.

iv) *Objective Evidence of Non-obviousness*

In addition to the lack of a prima facie case of obviousness, the strong objective evidence of non-obviousness presented during prosecution of this case further compels allowance of the claims.

Evidence of such objective indicia of non-obviousness, the so-called "secondary considerations" must be considered in all obviousness determinations. Stratoflex, Inc. v. Aeroquip Corp., 713 F.2d 1530, 1538-1539 (1983):

> Indeed, evidence of secondary consideration may often be the most probative and cogent evidence in the record. It may often establish that an invention appearing to have been obvious in light of the prior art was not. It is to be considered as part of all the evidence, not just when the decision-maker remains in doubt after reviewing the art.

See also Graham v. John Deere, 383 U.S. 1, 148 USPQ 459 (1966).

This standard set forth by the Federal Circuit applies not only during litigation of issued patents, but to a determination of patentability during ex parte prosecution as well. In re Sernaker, 702 F.2d. 989 217 USPQ 1, 7 (Fed. Cir. 1983). However, in the instant case, the Examiner is not properly considered evidence of "long felt need and failure of others" in maintaining the present § 103 rejection.

Specifically, Appellants' provided the Holler reference as indicating that the field longed to make the claimed fusion proteins but could not. Holler reported that phage display had not yet proven successful in making scTCRs despite what he saw as extensive structural similarity between antibodies and TCR V regions. Appellants' invention addressed this need and succeeded by providing soluble fusion molecules with a bacteriophage coat protein linked to the scTCR. The Holler reference is highly probative of the difficulties the field had in making these molecules and should be given substantial weight by Examiner Schwadron. MPEP 716.01(b).

The Examiner completely dismissed the Holler reference on grounds that **"Holler et al. simply not familiar with the prior art".** See Part D, above. That is no basis for disregarding the Holler's statement that the field wanted but failed to produce the claimed invention. Appellants are under no burden to provide evidence of Holler's knowledge of the art in order to have the reference considered as objective indicia of non-obviousness. See Stratoflex, Inc. v. Aeroquip Corp., 713 F.2d 1530, 1538-1539 (1983); In re Sernaker, 702 F.2d. 989 217 USPQ 1, 7 (Fed. Cir. 1983); and MPEP 716.01(b).

Moreover, Examiner Schwadron's citation of Weidanz et al. to support his disregard of Holler is clearly improper. That reference is not prior art and cannot serve as a basis for ignoring Holler or substantiating the obviousness rejection on appeal.

It is requested that the Board consider Holler as objective evidence that workers in the field wanted, but could not make, the claimed invention.

CONCLUSIONS

For the Examiner's prima facie case to stand, he has the burden of showing that: 1) The cited references disclose or suggest fusing a bacteriophage coat protein to a scTCR; 2) there is a settled role for the bacteriophage coat protein in making fusion proteins; and that 3) one could fuse Chung's scTCR to Barbas' coat protein with a reasonable expectation of success. These points have not been made by the Examiner. As discussed above, Barbas does not disclose scTCRs. Moreover, there was significant uncertainty in the field about whether it was possible to fuse a bacteriophage coat protein to a scTCR as exemplified by Huse and Onda. In view of the cited art and in consideration of the Examiner's position, it could be argued that one might be motivated to test fusing the bacteriophage coat protein to Chung's scTCR in the hope of producing a soluble and function protein. But this is not the legal standard required by our case law. It is without a doubt not obvious from the art of record to make the claimed invention of a scTCR fused to a bacteriophage coat protein.

Appellants submit that they have overcome the Examiner's obviousness rejection in the view of all the facts and argument of record in this case. Simply put, one of skill in this area would not be able to predict, with any reasonable expectation of success, how to make and use the claimed invention.

Importantly, Appellants have provided experimental evidence clearly showing that it is indeed possible to make and use scTCR-bacteriophage coat protein fusions. See Examples 1-11, 15 and 16 as discussed above.

In summary, Appellants submit that the instant invention is both novel and unobvious. The arguments set forth above establish that non-obviousness.

Although it is not believed that the present submission requires any fee for consideration by the Office, the Examiner is authorized to charge such fee to our deposit account **04-1105** should such fee be deemed necessary.

Respectfully submitted,

Date: August 5, 2004

By: _____

Robert L. Buchanan (Reg. No. 40,927)
EDWARDS & ANGELL, LLP
P. O. Box 55874
Boston, MA 02205
Tel. (617) 439-4444
Fax (617) 439-4170 / 7748
Customer No.: 21874

Doc. 453669

EXHIBIT A
Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72
on appeal

What is claimed is:

1. A soluble fusion protein comprising a bacteriophage coat protein covalently linked to a single-chain T cell receptor comprising an antigen binding pocket, wherein the single-chain T cell receptor comprises a V-α region covalently linked to a V-β region by a peptide linker sequence that effectively positions the V-α region and the V-β region to form the antigen binding pocket, the soluble fusion protein further comprising a C-β region fragment.

2. The soluble fusion protein of claim 1, wherein the C-terminus of the V-α region is covalently linked by the peptide linker sequence to the N-terminus of V-β region.

4. The soluble fusion protein of claim 2 wherein the C-β region fragment is covalently linked between the C-terminus of the V-β region and the N-terminus of the bacteriophage coat protein.

7. The soluble fusion protein of claim 2, wherein the peptide linker sequence contains from approximately 2 to 20 amino acids.

8. The soluble fusion protein of claim 1, wherein the bacteriophage coat protein is gene III or gene VIII protein.

14. A soluble fusion protein comprising covalently linked in sequence: 1) a V-α region, 2) a peptide linker sequence, 3) a V-β region covalently linked to a C-β region fragment, and 4) a bacteriophage gene VIII protein, wherein the peptide linker sequence effectively positions the V-α region and the V-β region to form an antigen binding pocket.

67. The soluble fusion protein of claim 1, wherein the C-terminus of the V-β region is covalently linked to the N-terminus of a C-β region fragment.

69. The soluble fusion protein of claim 1, wherein the V-α region and the V-β region are about 200 to 400 amino acids in length.

71. The soluble fusion protein of claim 1, wherein the C-β region fragment is about 50 to 126 amino acids in length.

72. The soluble fusion protein of claim 70, wherein the C-β region fragment does not include a cysteine residue corresponding to position 127 of a full-length C-β region.

#453669

0022-1767/92/14912-3914$02.00/0
THE JOURNAL OF IMMUNOLOGY

Vol. 149, 3914–3920, No. 12, December 15, 1992
Printed in U.S.A.

APPLICATION OF A FILAMENTOUS PHAGE pVIII FUSION PROTEIN SYSTEM SUITABLE FOR EFFICIENT PRODUCTION, SCREENING, AND MUTAGENESIS OF F(ab) ANTIBODY FRAGMENTS

WILLIAM D. HUSE,[1]* TIMOTHY J. STINCHCOMBE,* SCOTT M. GLASER,* LISA STARR,† MICHAEL MacLEAN,† KARL E. HELLSTRÖM,† INGEGERD HELLSTRÖM,† AND DALE E. YELTON*

*From *Ixsys, Inc., San Diego, CA 92121; †Department of Immunology, The Scripps Research Institute, La Jolla, CA 92037; and ‡Bristol-Myers Squibb Pharmaceutical Research Institute, Seattle, WA 98121*

We describe the application of a novel filamentous phage vector system suitable for efficient screening and production of F(ab) antibody fragments. The vector system can concurrently produce free F(ab) fragments and F(ab) displayed on the surface of M13 bacteriophage via a V_HC_H1-pVIII fusion protein. When expressed in a *supO* (nonsuppressor) strain of *Escherichia coli* free F(ab) can be produced. Antibody F(ab) fragments are secreted into culture medium at concentrations up to 0.3 mg/liter and conveniently subjected to detailed analysis with little or no purification. Higher concentrations of F(ab) (approximately 10 mg/liter) were found to accumulate in the periplasmic space. In this report the vector system is shown to produce correctly folded and assembled F(ab) fragments of chimeric L6, a mAb against a tumor-associated Ag expressed by many human carcinomas.

Until recently mAb have been primarily produced in mammalian cells. The slow growth rates and difficulty in genetically manipulating antibody genes expressed in mammalian cells have motivated development of methods to express antibody genes in simpler organisms. Molecular cloning techniques in bacteria have facilitated the production and manipulation of antibody fragments, increasingly aiding the search for useful antibodies (1, 2). Methods have been reported for the expression of antibody fragments in *Escherichia coli* using plasmids (3–7), bacteriophage λ (8–11), and more recently the filamentous phage M13 (12–17). Very large combinatorial libraries of 10^5 to 10^8 distinct antibody specificities can be created in microorganisms, far greater than can be achieved with hybridoma cell fusion methods (9, 10). However, the rapid identification, isolation, and, if necessary, modification of antibodies with the goal of improving affinity or redirecting specificity for Ag depends heavily on the availability of powerful screening methods both in terms of sampling large numbers of antibody fragments and evaluating certain aspects of antibody binding. Hence, an antibody expression vector that permits rapid cloning and mutagenesis of antibody V region genes and produces sufficient levels of antibody for biochemical analysis would be highly desirable.

Because antibody fragments can be displayed on the surface of filamentous phage, bacteriophage M13 vectors are proving particularly valuable in creating and screening sequence libraries for antibody fragments of interest. Briefly, fusion proteins are created by inserting DNA encoding an antibody fragment in front of a phage coat protein gene (18, 19). The fusion proteins become anchored in the phage coat via the coat protein and antibody sequence is displayed at the phage surface. Phage-bearing antibody sequences of interest can be detected by Ag binding and isolated in infectious form (13, 14). In addition to antibodies the incorporation of malarial protein (20), growth hormone (21), and a hexapeptide library (22) into the coat proteins of filamentous phage has been reported.

We have developed an M13 filamentous phage vector system that can produce and display F(ab) as a fusion product to pVIII coat protein and can also synthesize free F(ab) in quantity. We report here the production of chimeric L6 antibody F(ab) fragments in these M13-derived vectors. L6 is a mAb against a tumor-associated cell surface Ag expressed by many human carcinomas (23). L6 has been shown capable of lysing cancer cells in vitro (24) and is currently the subject of clinical trials (25).

MATERIALS AND METHODS

Construction of bacteriophage M13IXL604 for expression of L6 F(ab). Restriction enzymes, calf intestine alkaline phosphatase, T4 polynucleotide kinase, T4 DNA polymerase, and T4 DNA ligase were purchased from Boehringer-Mannheim (Indianapolis, IN). Total RNA was isolated from the chimeric L6-secreting cell line described by Fell et al. (26) by the guanidinium thiocyanate-phenol/chloroform method (27). First strand cDNA was synthesized using oligo dT and BRL Superscript reverse transcriptase (GIBCO BRL, Grand Island, NY) and PCR[2] amplification of L6 H chain (V_H-C_H1) and L chain (V_L-C_L) sequences was performed by the method of Saiki et al. (28) as modified by Sastry et al. (29). The following primers were used for PCR amplification. The restriction endonuclease recognizing the boldface and underlined cloning site in the sequences is indicated within the parentheses: forward V_H primer (*Xho*I): 5′-CAGTCTGGA-CCTGAG**CTCGAG**AAGCCTGGAGAG-3′; forward V_L primer (*Nco*I)

Received for publication June 17, 1992.
Accepted for publication September 25, 1992.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
[1] Address correspondence and reprint requests to Dr. William D. Huse, IXSYS, Inc., Suite L103, Bldg. 2, 3550 General Atomics Court, San Diego, CA 92121.

[2] Abbreviations used in this paper: PCR, polymerase chain reaction; gVIII, gene VIII coding for the major coat protein of M13 phage; pVIII, major coat protein of M13 phage; C_H1 H chain C region 1.; V_L, L chain V region; IPTG, isopropyl-β-D-thiogalactopyranoside; V_H, H chain V region; MOPS, 3-(*N*-morpholino)propane sulfonic acid; TES, 30 mM Tris-HCl, 2 mM EDTA, 2% sucrose (w/v), pH 8.0; *supE*, strains of *E. coli* that carry a glutamine inserting amber (UAG) suppressor tRNA; MES, 20 mM 3-(*N*-morpholino)propane sulfonic acid, 2 mM EDTA, 20% sucrose (w/v), pH 7.5; *supO*, strains of *E. coli* that do not carry a suppressor tRNA.

5′-GCCCAACCAGCCATGGCCCAAATTGTTCTCTCCCAGTCT-3′; reverse C_H1 primer (*Spe*I): 5′-TGTGTGAGTACTAGTACAAGATTT-GGGCTCTGC-3′; reverse C_κ primer (*Xba*I): 5′-CCGCTTAACTCT-AGACTAACACTCTCCCCTGTTCAAGCT-3′. The V_H-C_H1 and V_L-C_κ chains were digested and ligated into prepared M13IX31[3] and M13IX12 vectors (30) resulting in M13IX31L6-H and M13IX12L6-L, respectively (Fig. 1). Sequence corrections of the cloned L6 V_H-C_H1 and V_L-C_κ chains were accomplished by site-directed mutagenesis of uracil-substituted ssDNA as described (31, 32). The correcting nucleotides are indicated by the underlined sequences. Five N-terminal V_H amino acids were inadvertently omitted from the original L6 sequence information resulting in an 11-amino acid deletion upon cloning. One V_H correction primer was 5′-CTCTCCAGGCTTCTTC-AGCTCAGGTCCAGAGGCTTTTG(C/T)CAC-3′. This primer corrected the *Xho*I cloning site back to the original L6 sequence and replaced V_H codons 7 to 10. The C/T mixed site introduced an amino acid change in the leader sequence that was found to increase expression. The following V_H primer served to restore the remaining six N-terminal amino acids 5′-CAGCTCAGGTCCAGACTGCACCAAC-TGGATCTGGGCCATGGCTGGTTGGGC-3′. The following V_L primer served to correct an incorrect nucleotide in the reverse C_κ PCR primer: 5′-ACTCTCCCCTGTTGAAGCTCTTTGTGA-3′. The H chain encoding M13IX31L6-H and L chain encoding M13IX12L6-L vectors were combined by annealing as described (30) to form M13IXL604 (Fig. 1). M13IXL605 was derived from M13IXL604 by mutation of the TAG (stop) codon, located between the H chain encoding sequence and the pseudo wild-type gene VIII, to GGT (glycine) using the oligonucleotide 5′-CGCCTTCAGCACCGGATCCACTAGT-3′ so that continual V_H-C_H1-pVIII fusion protein would be made. DNA sequence analysis of ssDNA prepared from phage isolates was performed with Sequenase Version 2 according to the manufacturer (United States Biochemical, Cleveland, OH).

Antibodies and reagents. The anti-Id mAb to L6 have been previously described (33). Anti-Id 1 is a γ2b isotype, anti-Id 3 is a γ2a isotype, and both anti-Id 7 and anti-Id 13 are γ1 isotypes. Alkaline phosphatase-conjugated antibodies were purchased from Fisher Biotech (San Francisco, CA). Unconjugated goat anti-human κ antibody was purchased from Caltag Laboratories (So. San Francisco, CA). Rabbit anti-M13 IgG was purified by Sepharose-protein A chromatography and exhaustively absorbed against whole *E. coli*. The anti-M13 antibody was subsequently biotinylated using D-biotinoyl-ε-aminocaproic acid *N*-hydroxysuccinimide ester (Boehringer-Mannheim) using standard chemistries. Vectastain avidin-horseradish peroxidase complex (Vector Laboratories, Burlingame, CA) and streptavidin-alkaline phosphatase complex (Boehringer-Mannheim) were used for second step reactions.

Screening by replicate filter lifts. M13IXL604 phage were plated at a low plaque density. F(ab) expression was induced by overlaying the plate with a 0.45-μ nitrocellulose filter (Schleicher and Schuell, Keene, NH) soaked in 10 mM IPTG and incubating at room temperature from 6 h to overnight. The filter was removed and placed in blocking buffer (Biosite Diagnostics, San Diego, CA) to block nonspecific binding sites. Phage growth was resumed by incubating the plate for an additional 2 h at 37°C and a second filter applied as described above. This procedure was repeated for the last filter and all filters were then blocked in blocking buffer. All monoclonal or polyclonal antibodies to be used for screening were diluted in blocking buffer. Filters were probed with either alkaline phosphatase-conjugated goat antibody to human λ-chain, alkaline phosphatase-conjugated goat antibody to human κ-chain, or anti-Id 3, which binds to L6 antibody. In the case of anti-Id 3, a secondary alkaline phosphatase-conjugated goat antibody to mouse IgγZa was used for detection. All filters were then washed three times for 10 min with 25 mM Tris, 0.137 M NaCl, 5 mM KCl, 0.9 mM $CaCl_2$, 0.5 mM $MgCl_2$, and 0.05% Tween 20 (pH 7.4) and developed with alkaline phosphatase substrate reagent (Bio-Rad, Richmond, CA).

F(ab) production and purification. For analytic scale production of F(ab) the *supE* amber suppressor strain XL-1 (Stratagene, San Diego, CA) and the *supO* nonsuppressor strain MK30-3 (Boehringer-Mannheim) were each grown in 2X YT medium at 37°C until the cultures reached a density of 0.4 to 0.6 at OD_{600}. Each strain was then diluted 1/10 into three culture tubes containing 3 ml 2X YT and infected with 3 μl of high titer (10^{11} plaque-forming units/ml) phage stock of M13IXL604, M13IXL605, or M13IX31/tube and incubated with shaking for 3 h at 37°C. Protein synthesis was induced by the addition of IPTG to a final concentration of 1 mM and shaking allowed to proceed for 10 to 14 h at ambient temperature. Periplasmic fractions were prepared essentially as described by Skerra and Plückthorn (5). The infected cultures were centrifuged for 10 min at 4°C and the cleared culture supernatant was reserved. The bacterial pellet was resuspended in 60 ml cold TES and digested on ice for 10 min after adding an equal volume of a cold freshly prepared solution containing 2 mg/ml lysozyme in TES. The periplasmic space containing free F(ab) was fractionated by centrifugation at 9000 rpm for 10 min at 4°C. The soluble periplasmic fraction was retained and diluted in blocking buffer to adjust the concentration of F(ab) to be equivalent to that found in the cleared culture supernatant.

For purification of L6 F(ab) a 1-liter culture of MK30-3 was grown and infected with M13IXL604 as described above. The cells were harvested by centrifugation at 5800 × *g* for 10 min at 4°C. The pellet was resuspended in 20 ml MES at ambient temperature. An equal volume of 2 mg/ml lysozyme in MES was added with mild vortexing and the suspension incubated at ambient temperature for 10 min. The soluble periplasmic fraction was isolated by centrifugation at 9700 × *g* for 10 min at 4°C. The periplasmic fraction was subjected to a second centrifugation at 12,000 × *g* for 30 min at 4°C. The cleared periplasmic fraction was loaded onto a macroprep 50 S support (Bio-Rad) washed extensively with 20 mM MOPS, pH 7.5 and eluted with 20 mM MOPS, 120 mM NaCl, pH 7.5. The partially purified L6 F(ab) was concentrated by centrifugation in a Centricon 30 device (Amicon, Beverly, MA) and size fractionated on a Biogel P-60 (Bio-Rad) column equilibrated and eluted with 20 mM MOPS, 120 mM NaCl, pH 7.5. Eluted L6 F(ab) fractions were assayed for binding to anti-Id 3 by ELISA. Fractions eluting at 8, 9, and 10 ml were then pooled and concentrated. Samples from each purification step were analyzed by SDS-PAGE on a 10% nonreducing gel followed by staining with Coomassie brilliant blue.

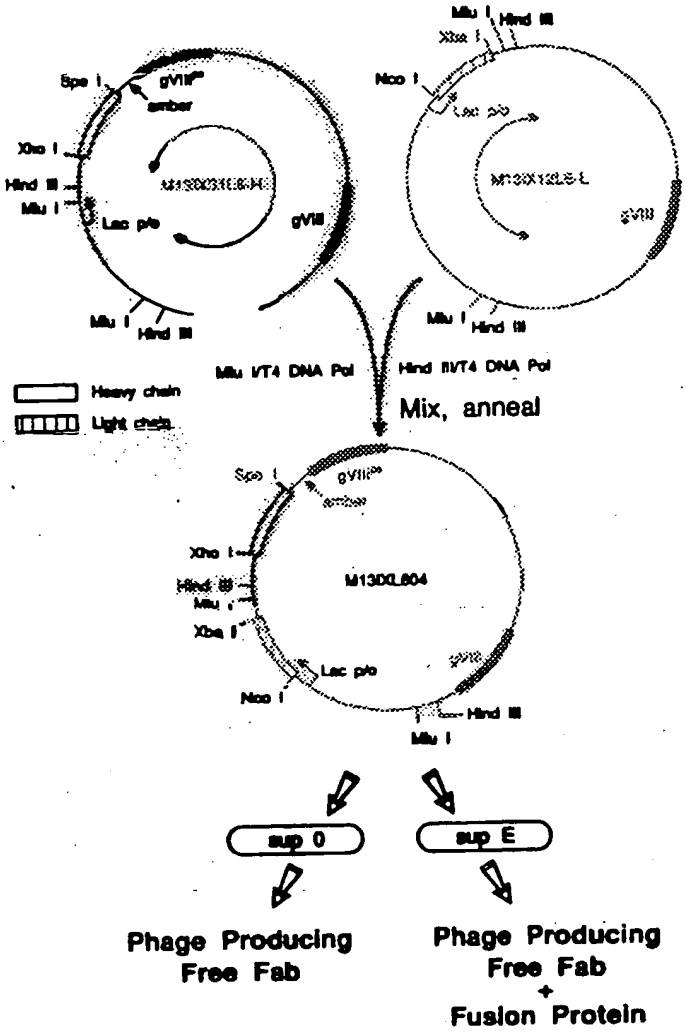


Figure 1. Construction of M13IXL604 vector for expression of L6 Fab. The M13IXL604 vector product contains one copy of pseudo wild-type gVIII (gVIII^{ps}) and one wild-type gVIII (gVIII) both downstream from the H chain encoding region. gVIII^{ps} has been altered in sequence to reduce homologous recombination with gVIII contained on the same vector. The presence of gVIII reduces the likelihood of selection against certain fusion proteins as the result of compromised phage viability (34).

[3] M13IX31 and M13IX12 vectors are available at no charge from Ixsys, San Diego, CA.

ELISA characterization of L6 F(ab) produced by M13IXL604. f13IXL604 phage stocks were prepared for assay by overnight nfection of XL-1 in LB broth and 10 μg/ml tetracycline to maintain xpression of F', and 1 mM IPTG. The bacteria were removed by entrifugation (10 min in a microfuge) and the phage solution was ssayed for binding to L6-specific anti-Id 1, 3, 7, 13, and to an anti--chain antibody by ELISA. All antibodies were coated onto Immulon microtiter plates (Dynatech Laboratories, Chantilly, VA) at 10 μg/ ml in 0.1 M $NaHCO_3$, pH 8.5, overnight at 4°C. The antibody solution vas shaken out and the plates were blocked with 300 μl of specimen locking buffer (Genetic Systems, Seattle, WA) for 1 h at room emperature. Plates were washed before use with 0.5% Tween 20 in .15 M NaCl and with the same solution between incubations. Phage ilutions were dispensed into microtiter plates in 100 μl of blocking uffer and incubated overnight. After washing, 100 μl of conjugate locking buffer (Genetic Systems) containing 1 μg/ml of biotinylated abbit antibodies to M13 were incubated 1 h at ambient temperature. fter washing, 100 μl of Vectastain avidin-HRP complex were incubated 30 min at room temperature. After a final wash step, 100 l of chromogenic substrate (3,3',3,5'-tetramethylbenzidine) in a itrate/phosphate buffer were added. The reaction was stopped with 00 μl of 3 N H_2SO_4 at various times ranging from 10 min to 2 h to chieve an optimum signal to background ratio for each different late-coating antibody. Plates were read on a microplate reader Biotek, Burlington, VT) in dual channel mode at 450/630 nm. All ssay points were measured in duplicate. M13IX31 served as a egative control.

The expression of M13IXL604, M13IXL605, and M13IX31 in acterial strains MK30–3 and XL-1 was also analyzed by ELISA. ulture supernatant and diluted periplasmic fractions prepared from nfected XL-1 and MK30–3 bacteria described above were serially iluted in diluent (Biosite Diagnostics), added to microtiter plates oated with anti-Id 3 and incubated for 2 h at ambient temperature. 6 F(ab) expressed in the culture medium or in the periplasmic space vas detected with an alkaline phosphatase-conjugated goat antibody o human κ-chains. Expression of L6 F(ab) displayed on the phage urface was detected by incubating the captured sample first with he biotinylated rabbit antibody to M13 and then with streptavidin-lkaline phosphatase complex. The washed plates were developed vith 6 mg/ml phenolphthalein monophosphate in 0.1 M aminomethylpropanediol, 0.5 M Tris, and 0.1% NaN_3, pH 10.2 (JBL Scientific, an Luis Obispo, CA) for 10 to 30 min. The reaction was stopped by ddition of one-third volume of cold 30 mM Tris base, 6 mM EDTA, nd the absorbance at 560 nm measured.

RESULTS

Construction of L6 F(ab) expression vector f13IXL604. The construction of M13IXL604 is shown n Figure 1. Total RNA was isolated from the chimeric .6-secreting cell line and cDNA synthesized. After PCR mplification of the cDNA using sequence-specific rimers, the chimeric L6 V_L-Cκ L chain was digested to ompletion with *NcoI* and *XbaI* and cloned into the *NcoI*/ *XbaI* site of M13IX12 to construct M13IX12L6-L. Similarly, the chimeric L6 V_H-C_H1 H chain PCR product was igested to completion with *XhoI* and *SpeI* and cloned nto the *XhoI*/*SpeI* site of M13IX31 to construct f13IX31L6-H. M13IX31 contains an amber stop codon ocated directly 5' of a modified gVIII gene (pseudo-gVIII) ncoding mature M13 pVIII major coat protein. Cloning ntibody V_H or V_H-C_H1 regions into the *XhoI*/*SpeI* site of f13IX31 abuts these antibody sequences in frame with he amber stop-pseudo gVIII sequence. Thus, f13IX31L6-H should produce L6 V_H-C_H1-pVIII fusion roduct when expressed in an *E. coli* amber suppressor train such as XL-1 and produce predominantly free L6 $_H$-C_H1 protein when expressed in a nonsuppressor train such as MK30–3. The H and L chain vectors were hen recombined into the single M13IXL604 expression hage by annealing *Hin*dIII/T4 DNA polymerase-digested f13IX12L6-L to *Mlu*I/T4 DNA polymerase-digested f13IX31L6-H through the homologous regions found between the *Hin*dIII and *Mlu*I restriction sites contained in oth vectors (30). Correct L6 H and L chain sequences were confirmed by DNA sequence analysis. The vector M13IXL605, which contains a GGT codon coding for glycine in place of the amber stop found in M13IXL604, was constructed by site-directed mutagenesis of M13IXL604. Thus, M13IXL605 should produce L6 V_H-C_H1-pVIII fusion product irrespective of the bacterial host used for expression.

Functional characterization and purification of M13IXL604-expressed L6 F(ab). M13IXL604 phage expressing L6 F(ab) was initially characterized by replicate filter lift assays (Fig. 2). XL-1 bacteria were infected at a low multiplicity of infection resulting in low plaque density of phage. No positive plaques were detected by filter lift assay when probed with alkaline phosphatase-conjugated goat antibody to human λL chain (Fig. 2*A*). In contrast, numerous, superimposable, uniform positive plaques were detected in assays of human κ-chain (Fig. 2*B*) and anti-Id 3, which recognizes assembled L6 H and L chain (Fig. 2*C*). Thus, screening by replicate filter lifts allows for detection of functional F(ab) expressed in this system.

To further evaluate the functional integrity of L6 F(ab) expressed by M13IXL604, phage displaying L6 F(ab) on the phage surface were assayed for binding to a panel of distinct mouse monoclonal anti-Id raised against the mouse mAb L6. ELISA experiments shown in Figure 3 demonstrated that four anti-Id antibodies specific for L6

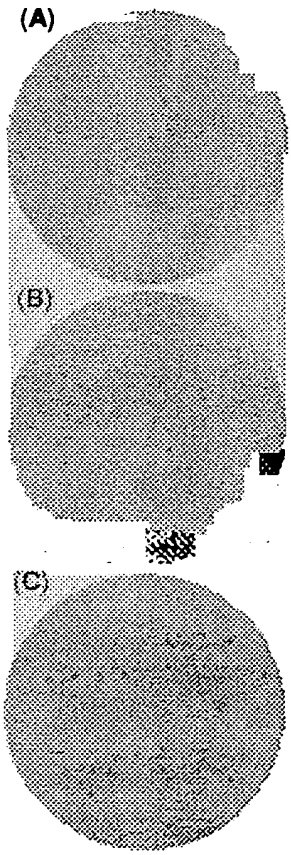


Figure 2. Screening for L6 F(ab) reactivity by replicate filter lifts. M13IXL604 phage was replated at low density for screening. Replicate filter lifts were prepared and probed with *A*) alkaline phosphatase-conjugated goat antibody to human λ-chain, *B*) alkaline phosphatase-conjugated goat antibody to human κ-chain, and *C*) anti-Id 3 antibody. The filter in C was washed and incubated with alkaline phosphatase-conjugated goat antibody to mouse Igγ2a. Filters were then washed and developed with alkaline phosphatase substrate reagent.

are recognized by the fusion protein displayed on the M13IXL604 phage. In the ELISA format used, M13IXL604 phage were grown in the suppressor bacterial strain XL-1, and phage displaying incorporated L6 F(ab)-pVIII protein are captured by solid phase anti-Id antibodies and detected with anti-M13 antibodies. Thus, only phage-bound L6 F(ab) is detected. The four anti-Id antibodies exhibit different binding specificities to L6 (33). Anti-Id 1 and 7 bind to L6 L chain, 13 to H chain and 3 to assembled L and H chains. The binding of M13IXL604 phage to an anti-κ-chain antibody further confirmed the assembly of L6 L chain with L6 H chain-pVIII fusion protein to yield F(ab), because L6 L chains are of the κ type (Fig. 3). M13IX31, which contains no H chain-encoding region served as a negative control for each anti-Id. A representative curve shows no M13IX31 binding. The ELISA results in Figure 3 thus indicate the presence of both L6 L and H chains and confirm that proper folding and assembly of L and H chains on the phage surface have been achieved.

Purification of L6 F(ab) fragments produced by M13IXL604 phage grown in the nonsuppressor bacterial strain MK30-3 was accomplished by loading the crude periplasmic fraction prepared from a 1-liter shake flask culture onto a cation exchange resin. The eluted material was concentrated and size fractionated by gel exclusion chromatography. Figure 4*A* shows the elution profile from the gel exclusion column as assayed by ELISA for functional binding to anti-Id 3 using anti-human κ to detect captured L6 F(ab). The purified F(ab) was shown to have a M_r of 42.7 kDa by SDS-PAGE, a value consistent with the 48-kDa size of known F(ab) fragments (Fig. 4*B*). Based therefore on both functional and physical properties, the identity of the purified material is chimeric L6 F(ab).

Regulated expression of M13IXL604 production of F(ab) and phage displayed F(ab). The proper activity of M13IXL604-expressed L6 F(ab) having been established, the functioning of the expression control system of M13IXL604 was tested. This involved comparing levels of fusion F(ab) with free F(ab) produced when phage were grown in suppressor strains (*supE*) vs nonsuppressor

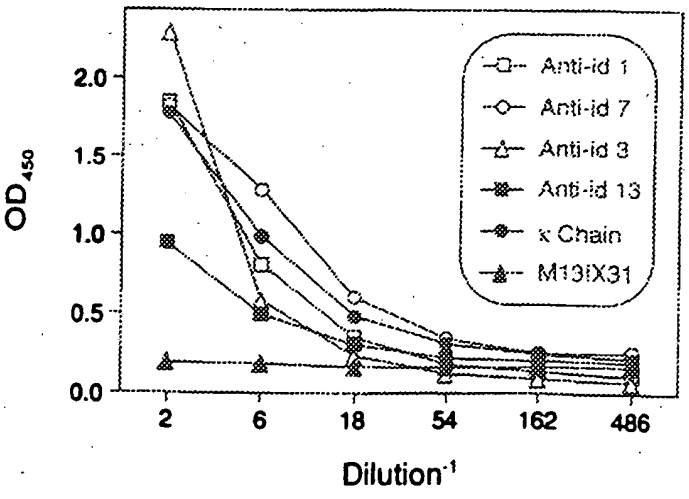
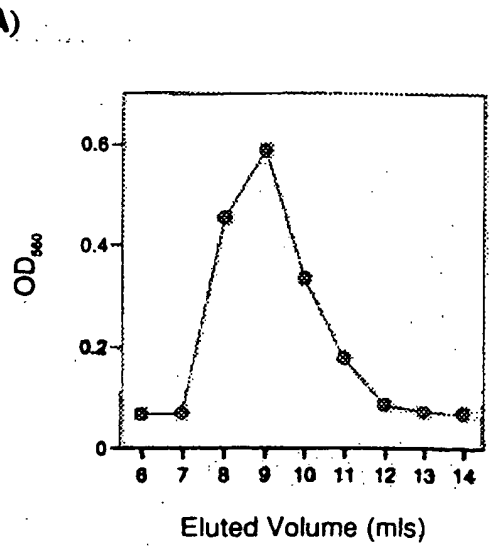


Figure 3. ELISA characterization of L6 F(ab) produced by M13IXL604. Various dilutions of phage expressing L6 F(ab) on the phage surface were evaluated for binding to anti-Id 1, 3, 7, and 13 and to an anti-human κ-chain antibody. Bound phage were then detected as described in *Materials and Methods*. Curves represent the average of triplicate assays.



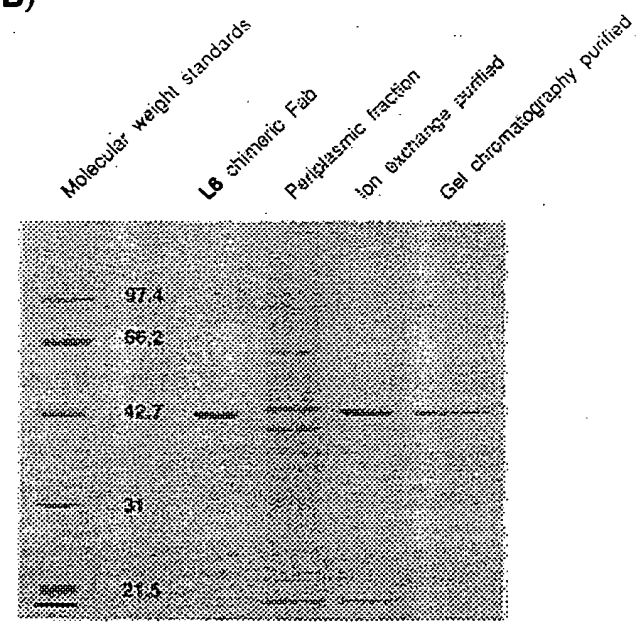


Figure 4. Purification of L6 F(ab) expressed by M13IXL604-infected MK30-3 bacteria. *A*) elution profile of L6 F(ab) fractions assayed for binding to anti-Id 3 antibody by ELISA. Fractions eluting at 8, 9, and 10 ml were then pooled and concentrated. *B*) SDS-PAGE analysis of purified Fab. Samples from each purification step were analyzed by SDS-PAGE on a 10% nonreducing gel followed by staining with Coomassie brilliant blue. The L6 chimeric F(ab) standard was prepared from whole IgG by papain digestion followed by protein A chromatography (Pierce, Rockford, IL).

strains (*supO*) of bacteria, respectively. This comparison was made using ELISA assays employing anti-Id 3 to detect assembled L6 F(ab). F(ab) displayed on the surface of M13 phage was detected using the M13 assay described in Figure 2. M13IXL605, a variant construct of M13IXL604 that has the amber stop codon replaced with a glycine encoding codon and therefore always synthesizes V_H-C_H1-pVIII protein, served as a control.

When phage were grown in the *supE* strain XL-1, both M13IXL604 and M13IXL605 secreted phage displaying F(ab) into the culture medium (Fig. 5*A*) and, as expected, little or no mature phage displaying F(ab) was detected in the periplasmic space (Fig. 5*B*). In the *supO* strain

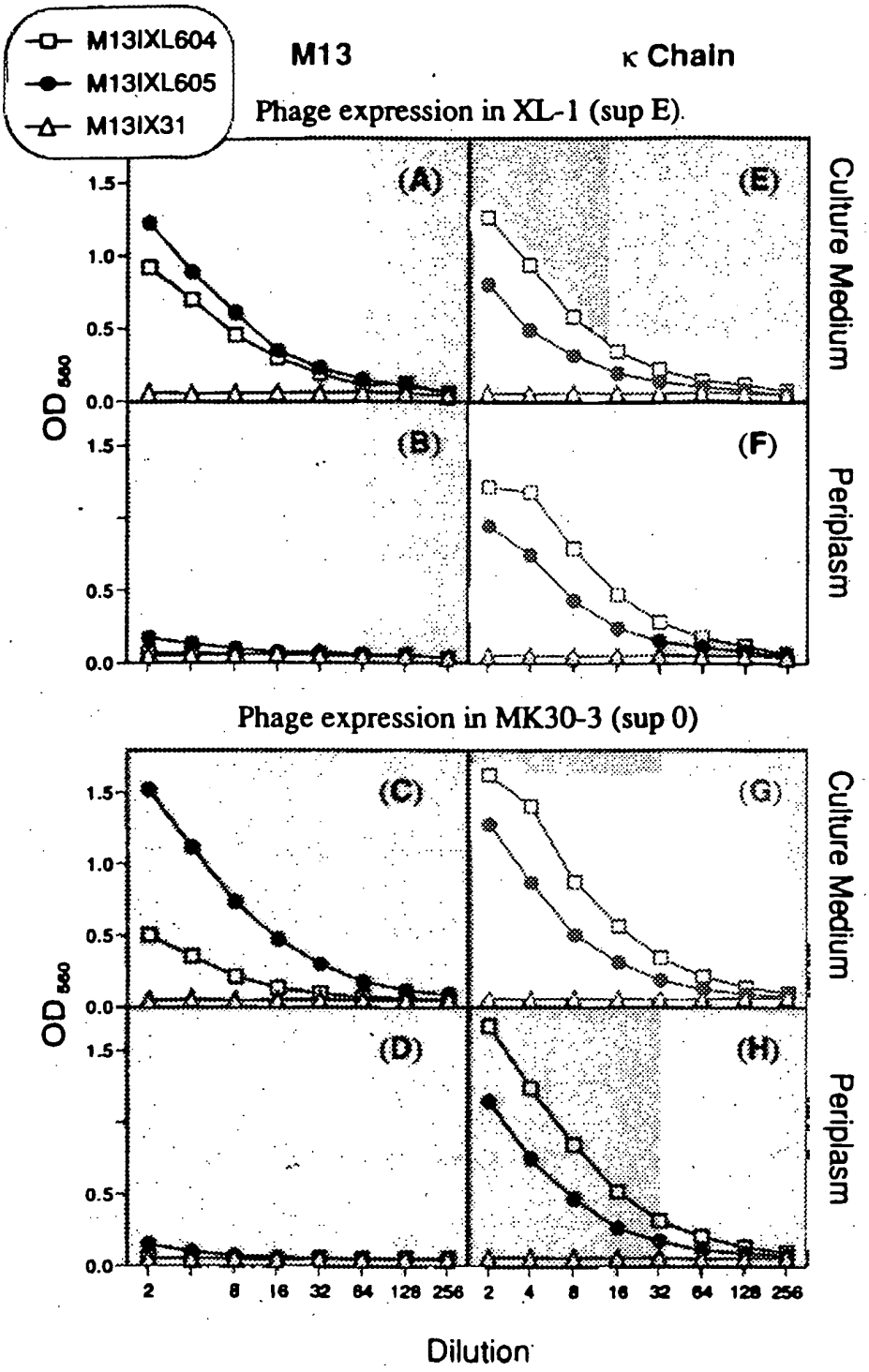


Figure 5. Characterization of the 3IXL604 expression control system. uid cultures of the amber suppres- strain XL-1 (*supE*) and the non- pressor strain MK30–3 (*supO*) were h infected with either M13IXL604, 3IXL605, or M13IX31 and culture su- natant and periplasmic fractions pre- ed as described in *Materials and hods*. Culture supernatant and peri- sm from infected cultures were as- ed for functional binding to anti-Id 3 F(ab) displayed on phage surface (*A* to and by total F(ab) (*E* to *H*).

30–3, the amber stop codon in M13IXL604 should vent fusion protein formation. Figure 5*C* shows phage playing F(ab) to be sharply reduced in M13IXL604 vs .3IXL605. The low signal seen in the M13IXL604 cul- e medium is presumably caused by nonspecific asso- tion of L6 F(ab) with phage particles and/or transla- nal readthrough. Again no mature phage displaying b) was found in the periplasm (Fig. 5*D*). In examining production of F(ab) by the anti-Id 3/κ-chain assay h M13IXL604 and M13IXL605 secreted anti-κ-reac- e material into the culture medium when expressed in her XL-1 or MK30–3 (Figs. 5*E* and 5*G*). Because the efficiency of amber suppression can be widely variable, we expect that M13IXL604 expressed in XL-1 should secrete free F(ab) in addition to F(ab) displayed on the surface of phage. Thus, the signal in Figure 5*E* should represent some combination of the two molecules. Polyethylene glycol precipitation of both M13IXL604 and M13IXL605 phage from induced XL-1 culture supernatants showed that approximately 95% of the anti-κ-reactive material remained in the polyethylene glycol-cleared supernatant (data not shown) indicating that the predominant secreted product is free L6 F(ab). Consistent with the view that M13IXL604 and M13IXL605 secrete free

L6 F(ab), abundant amounts of soluble assembled F(ab) were detected in the periplasmic space in both strains of *E. coli* (Fig. 5, *F* and *H*). MK30–3 appeared to consistently produce higher concentrations of F(ab) than XL-1. Quantitative ELISA demonstrated that the concentration of free F(ab) produced by M13IXL604 grown in the *supO* strain MK30–3 approaches values up to 0.3 mg/liter in culture supernatant and concentrations approaching 10 mg/liter within the periplasmic space (data not shown).

DISCUSSION

M13IX31 (for H chains) and M13IX12 (for L chains) are the M13 vectors that can receive polyclonal sets of V regions for constructing combinatorial libraries or individual antibody genes for V region production and/or mutagenesis (30). The system is the first to allow differentially controllable production of free and fusion F(ab). Upon recombining the two vectors, all of the control elements required for regulated expression of phage displayed F(ab) or free F(ab) are contained in a single expression vector. For phage display F(ab) is synthesized as a fusion product to pVIII coat protein, the major structural protein of the filamentous M13 phage particle. In the M13 bacteriophage the major coat protein pVIII is expressed in several thousand copies per phage particle. Although this level of expression may be useful for producing F(ab) as F(ab)-pVIII fusion proteins, pVIII synthesized solely as a fusion product is not likely to form a properly assembled coat as a result of steric hindrance by the much larger H chain polypeptide. We reasoned that if wild-type pVIII, in addition to F(ab)-pVIII fusion product, is available for assembly then the formation of a mature infectious phage particle could occur. With this in mind, the H chain-gVIII expression vector M13IX31 was constructed to contain two copies of gVIII. The copy that anchors F(ab) to the surface of M13 has had gVIII codons extensively altered (pseudo gVIII) to prevent recombination with the wild-type gVIII.

To demonstrate the utility of these vectors we cloned the antibody V_L-Cκ- and V_H-C_H1-encoding sequences from the chimeric L6 transfectoma cell line into the M13IX12 and M13IX31 vectors, respectively. The two vectors were recombined to create M13IXL604. M13IXL604 expresses a dicistronic message encoding both H and L chain L6 sequences under transcriptional control of an inducible Lac promoter. An amber stop codon resides between the H chain-encoding region and the pseudo gVIII coat protein. When grown in an amber suppressor strain (*supE*) of *E. coli*, M13IXL604 was shown to produce L6 H chain-pVIII fusion protein resulting in phage-displayed F(ab) in addition to free F(ab). When free F(ab) was desired as the major end product, the phage was grown in a nonsuppressor strain (*supO*). The M13IXL605 vector, which solely produces F(ab)-pVIII fusion protein, was constructed to evaluate, by comparison, the M13IXL604 control system.

M13IXL605 was found to display slightly higher levels of phage-associated F(ab) than M13IXL604 when grown in the *supE* bacterial strain. This is likely attributed to a higher proportion of F(ab)-pVIII fusion protein relative to wild-type pVIII produced in the M13IXL605-infected strain. In both the *supE* and *supO* bacterial strains M13IXL604 secreted chimeric L6 F(ab) at levels somewhat higher than those of M13IXL605. Phage titers of M13IX31-, M13IXL604-, and M13IXL605-infected cultures were found to decrease relative to the level of F(ab)-pVIII fusion protein incorporation (D. Yelton, unpublished observations). Taken together, these results suggest that a functionally viable phage particle may be able to tolerate a limited number of incorporated F(ab)-pVIII fusion products and that the amount of F(ab) incorporated into the phage coat may inversely affect phage titers and overall F(ab) yield. Furthermore, the appearance of free F(ab) in the culture media suggests that H chain-pVIII fusion protein unincorporated into phage is properly assembled with L chain and secreted as functional F(ab).

The set of anti-Id antibodies raised against the murine mAb L6 demonstrated that faithful expression and function of bacterially produced chimeric L6 F(ab) occurs in the M13IX12 and M13IX31 vector system. L6 F(ab) activity is readily detectable by either the nitrocellulose filter lift or ELISA formats. The anti-Id antibodies used in these studies serve as a convenient and informative model for antibody-Ag binding, inasmuch as the tumor Ag bound by L6 in vivo is yet to be purified or fully characterized.

The M13IX31 and M13IX12 vector system described here can serve as a versatile, general purpose approach to F(ab) production and screening. F(ab) production in this system is sufficiently robust to permit multiple replicate filter lifts, a practical requirement for implementing various screening strategies. The vectors can be used to create combinatorial antibody libraries to identify novel antibodies. Screening of sufficiently large combinatorial antibody libraries could potentially allow useful antibody fragments of murine or human origin to be isolated without the necessity to perform standard immunization procedures (35). In addition, because oligonucleotide-directed mutagenesis is convenient and highly efficient in M13 these vectors are ideal for engineering antibodies with new properties. It is conceivable that the ability to create stable, high affinity human antibodies is a promising endeavor.

In a companion publication (36), we show that the M13IX31 and M13IX12 antibody expression system lends itself to efficient antibody engineering by site-directed mutagenesis. Codon-based mutagenesis of L6 hypervariable regions proved effective in altering the fine specificity of L6 in a predefined manner. Thus the system allows mutagenesis, screening, and mutant F(ab) production to be accomplished very rapidly.

Acknowledgments. We thank G. Valkirs and members of Biosite Diagnostics for useful discussions and for supplying the formulation for reagents used in this study. We thank S. L. Morrison for review of the manuscript and G. A. Cruz for technical assistance.

REFERENCES

1. Plückthorn, A. 1990. Antibodies from *Escherichia coli*. *Nature 347:497.*
2. Winter, G., and C. Milstein. 1991. Man-made antibodies. *Nature 349:293.*
3. Boss, M. A., J. H. Kenten, C. R. Wood, and J. S. Emtage. 1984. Assembly for functional antibodies from immunoglobulin heavy and light chains synthesized in *E. coli*. *Nucleic Acids Res. 12:3791.*
4. Cabilly, S., A. D. Riggs, H. Pande, J. E. Shively, W. E. Holmes, M. Rey, L. J. Perry, R. Wetzel, and H. L. Heyneker. 1984. Generation of antibody activity from immunoglobulin chains produced in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA 81:3273.*
5. Skerra, A., and A. Plückthorn. 1988. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science 240:1038.*

6. Better, M., C. P. Chang, R. R. Robinson, and A. H. Horwitz. 1988. Escherichia coli secretion of an active chimeric antibody fragment. *Science 240:1041.*
7. Ward, E. S., D. Güssow, A. D. Griffiths, P. T. Jones, and G. Winter. 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli. Nature 341:544.*
8. Sastry, L., M. Alting-Mees, W. D. Huse, J. M. Short, J. A. Sorge, B. N. Hay, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1989. Cloning of the immunological repertoire in *Escherichia coli* for generation of monoclonal catalytic antibodies: construction of a heavy chain variable region-specific cDNA library. *Proc. Natl. Acad. Sci. USA 86:5728.*
9. Huse, W. D., L. Sastry, S. A. Iverson, A. S. Kang, M. Alting-Mees, D. R. Burton, S. J. Benkovic, and R. A. Lerner. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science 246:1275.*
10. Caton, A. J., and H. Koprowski. 1990. Influenza virus hemagglutinin-specific antibodies isolated from a combinatorial expression library are closely related to the immune response of the donor. *Proc. Natl. Acad. Sci. USA 87:6450.*
11. Persson, M. A. A., R. H. Caothien, and D. R. Burton. 1991. Generation of diverse high-affinity human monoclonal antibodies by repertoire cloning. *Proc. Natl. Acad. Sci. USA 88:2432.*
12. McCafferty, J., A. D. Griffiths, G. Winter, and D. J. Chiswell. 1990. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature 348:552.*
13. Kang, A. S., C. F. Barbas, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1991. Linkage of recognition and replication functions by assembling combinatorial antibody F(ab) libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA 88:4363.*
14. Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter. 1991. Making antibody fragments using phage display libraries. *Nature 352:624.*
16. Breitling, F., S. Dubel, T. Seehaus, I. Klewinghaus, and M. Little. 1991. A surface expression vector for antibody screening. *Gene 104:147.*
17. Chang, C. N., N. F. Landolfi, and C. Queen. 1991. Expression of antibody Fab domains on bacteriophage surfaces: potential use for antibody selection. *J. Immunol. 147:3010.*
18. Smith, G. P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science 228:1315.*
19. Parmley, S. F., and G. P. Smith. 1988. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene 73:305.*
20. de la Cruz, V. F., A. A. Lal, and T. F. McCutchan. 1988. Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *J. Biol. Chem. 263:4318.*
21. Bass, S., R. Greene, and J. A. Wells. 1990. Hormone phage: an improved method for enrichment of variant proteins with altered binding properties. *Proteins 8:309.*
22. Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower. 1990. Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. USA 87:6378.*
23. Hellström, I., D. Horne, P. Linsley, J. P. Brown, V. Brankovan, and K. E. Hellström. 1986. Monoclonal mouse antibodies raised against human lung carcinoma. *Cancer Res. 46:3917.*
24. Hellström, I., P. L. Beaumier, and K. E. Hellström. 1986. Antitumor effects of L6, an IgG2a antibody that reacts with most human carcinomas. *Proc. Natl. Acad. Sci. USA 83:7059.*
25. Goodman, G. E., I. Hellström, L. Brodzinsky, C. Nicaise, B. Kulander, D. Hummel, and K. E. Hellström. 1990. Phase I trial of murine monoclonal antibody L6 in breast, colon, ovarian, and lung cancer. *J. Clin. Oncol. 8:1083.*
26. Fell, H. P., M. A. Gayle, D. Yelton, L. Lipsich, G. L. Schieven, J. S. Marken, A. Aruffo, K. E. Hellström, I. Hellström, and J. Bajorath. 1992. Chimeric L6 anti-tumor antibody: genomic construction, expression, and characterization of the antigen binding site. *J. Biol. Chem. In press.*
27. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem. 162:156.*
28. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science 239:487.*
29. Sastry, L., M. Alting-Meese, W. D. Huse, J. M. Short, J. A. Sorge, B. N. Hay, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1989. Cloning of the immunological repertoire in *Escherichia coli* for generation of monoclonal catalytic antibodies: construction of a heavy chain variable region-specific cDNA library. *Proc. Natl. Acad. Sci. USA 86:5728.*
30. Huse, W. D. Combinatorial antibody expression libraries in filamentous phage. In *Antibody Engineering: A Practical Guide,* C. A. K. Borrebaeck, ed. W. H. Freeman and Company, New York, 1991, pp. 103–120.
31. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA 82:488.*
32. Kunkel, T. A., J. Roberts, and R. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol. 154:367.*
33. Hellström, K. E., D. E. Yelton, H. P. Fell, D. Beaton, M. Gayle, M. MacLean, M. Kahn, and I. Hellström. 1990. Epitope mapping and use of anti-idiotype antibodies to the L6 monoclonal anticarcinoma antibody. *Cancer Res. 50:2449.*
34. Russel, M., and P. Model. 1981. A mutation downstream from the signal peptidase cleavage site affects cleavage but not membrane insertion of phage coat protein. *Proc. Natl. Acad. Sci. USA 78:1717.*
35. Marks, J. D., H. R. Hoogenboom, T. P. Bonnert, J. McCafferty, A. D. Griffiths, and G. Winter. 1991. By-passing immunization: human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol. 222:581.*
36. Glaser, S. M., D. E. Yelton, and W. D. Huse. 1992. Antibody engineering by codon-based mutagenesis in a filamentous phage vector system. *J. Immunol. 149:3903.*

Pergamon
0161-5890(95)00098-4

Molecular Immunology, Vol. 32, No. 17/18, pp. 1387–1397, 1995

Printed in Great Britain
0161-5890/95 $9.50 + 0.00

A PHAGE DISPLAY SYSTEM FOR DETECTION OF T CELL RECEPTOR–ANTIGEN INTERACTIONS

TAKEBUMI ONDA,*†‡ DRAKE LAFACE,*‡§ GOTTFRIED BAIER,*‖ THOMAS BRUNNER,* NAKAYUKI HONMA,¶ TOSHIFUMI MIKAYAMA,¶ AMNON ALTMAN* and DOUGLAS R. GREEN*

*Division of Cellular Immunology, La Jolla Institute for Allergy and Immunology, 11149 N. Torrey Pines Rd, La Jolla, CA 92037, U.S.A.; ¶Kirin Brewery Co. Ltd, Pharmaceutical Research Laboratory, Maebashi Shi Gunma; Japan

(*First received* 29 *June* 1995; *accepted in revised form* 30 *June* 1995)

Abstract—The process of T cell recognition involves a complex set of interactions between the various components of the TCR/MHC peptide trimolecular complex. We have developed a system for exploring the specific binding interactions contributed by the constituent subunits of TCR complexes for components of their ligands. We utilized an M13 phage display system, designed for multivalent receptor display, to explore specific binding interactions between various TCRα chains and specific antigen in the absence of MHC. The multivalent TCR–phage display system was sensitive enough to reveal some TCRα chains capable of binding directly to antigen with the same fine specificity shown by the MHC-restricted T cells from which the α chains were derived. Cross-specificity analysis using two antigen-binding TCRα chains derived from T cells with different polypeptide antigen specificities confirmed the fidelity of this binding. In mixtures of antigen-binding and non-binding TCRα-displaying phage, specific selection was achieved at a starting frequency of 1/1000, suggesting that this system can be employed for selection and analysis of TCR-displaying phage libraries. While the binding specificities exhibited by these TCRs are unusual, they provide a novel perspective from which to study the specific binding interactions that constitute TCR antigen binding.

Key words: M13 filamentous phage, peptide binding, phage-display, TCR.

INTRODUCTION

Specific recognition by T lymphocytes is mediated by T cell antigen receptors (TCR) (Jorgensen *et al.*, 1992b; Chien and Davis, 1993). The most extensively studied TCR–ligand interactions involve TCRαβ heterodimers binding to peptide–MHC complexes. A number of studies have focused on defining the intricate biochemical interactions between the various components of the TCR and MHC–peptide complexes. As a consequence of these investigations, it has been suggested that residues within the CDR3 regions of TCRα and β chains interact with specific residues from the antigenic peptide fragment (Davis and Bjorkman, 1988; Engel and Hedrick, 1988; Danska *et al.*, 1990; Jorgensen *et al.*, 1992a). More recent studies have indicated that TCR interactions with peptide appear to be very specific, while interactions with MHC are comparatively degenerate (Ehrich *et al.*, 1993). The TCR–MHC interactions could be established in a variety of configurations with the same TCR and MHC and appear to be influenced by the interaction of the TCR with the peptide. Thus, while TCR interactions with residues from both peptide (Engel and Hedrick, 1988; Danska *et al.*, 1990; Jorgensen *et al.*, 1992a) and MHC (Ajitkumar *et al.*, 1988; Peccoud *et al.*, 1990) appear to be essential for efficient T cell recognition, the dominant interactions are seemingly mediated by CDR3 residues associated with specific peptide (Ehrich *et al.*, 1993).

Although most studies have focused on conventional TCR binding of MHC–peptide ligands, a number of alternative TCR–ligand interactions have emerged, providing the basis for further studies in TCR specificity. Examples include TCR β chain-binding to superantigens (White *et al.*, 1989; Gascoigne and Ames, 1991; White *et al.*, 1993); TCRs which specifically bind peptide-free mycolic acid antigens in the context of CD1 restricting elements (Porcelli *et al.*, 1992); and TCRs with specificities for carbohydrate moieties of post-translationally modified peptide antigens (Haurum *et al.*, 1994; Michaëlsson *et al.*, 1994). These recent studies of non-conventional TCR–ligand interactions lend credence to earlier studies of T cells which expressed αβ TCRs capable of binding directly to hapten-conjugates (Rao *et al.*,

†Current address: National Children's Medical Research Center, Tokyo, Japan.
‡These authors contributed equally and share first authorship.
§Author to whom correspondence should be addressed: Drake M. LaFace, La Jolla Institute for Allergy and Immunology, 11149 N. Torrey Pines Rd, La Jolla, CA 92037, U.S.A.
‖Current address: Institute for Medical Biology and Human Genetics, University of Innsbruck, Innsbruck, Austria.

1984a, 1984b; Siliciano *et al.*, 1986) or peptide antigens (Esch and Thomas, 1990) in the absence of MHC proteins. Interestingly, direct TCR-binding interactions with hapten or peptide ligands, in the absence of MHC, demonstrated relatively high affinity constants (K_Ds = 5×10^{-5}–6×10^{-6}) when compared to the normal range of TCR–ligand affinities (Rao *et al.*, 1984b; Siliciano *et al.*, 1986; Esch and Thomas, 1990; Davis and Chien, 1993).

Studies of TCR that directly bind nitrophenol haptens suggested that the specificity was mediated by the TCRα subunit in a manner independent of TCRβ (Kuchroo *et al.*, 1991). Similar studies by our group and others have suggested the possibility that a subset of TCRα molecules may bind directly to antigen for which the T cell encoding the TCR protein is specific (Bissonnette *et al.*, 1991; Green *et al.*, 1991; Mori *et al.*, 1993). While these unusual TCR–ligand interactions may represent a relatively small subset of TCR specificities, they are nonetheless intriguing because they provide an alternative approach with which to explore the many complex interactions that mediate specific binding. Since the constituent α and β chains of TCR may bind different residues of the peptide ligand component in a relatively autonomous manner (Jorgensen *et al.*, 1992a, 1992b), we reasoned that it should be feasible to study directly the specific interactions of TCRα chains with antigenic peptide. An elemental analysis of the specific interactions that constitute the complex binding of TCR with MHC–peptide ligands would be extremely useful. While these studies would not refute existing theories of conventional TCR-antigen binding, they would provide a unique perspective from which to study the individual specific binding interactions between the various subunits of the TCR–MHC/peptide complex which contribute to the overall TCR binding specificity.

We employed the filamentous phage display system for the study of immunological receptors (Kang *et al.*, 1991; Barbas and Lerner, 1991b) to directly assess specific binding interactions of several TCRα polypeptides with antigenic peptide and globular proteins in the absence of MHC. The multivalent phage display system, utilized to enhance avidity, was optimally suited to study interactions of TCR molecules directly with their peptide ligand, as the binding affinities are relatively low compared to immunoglobulin receptors.

EXPERIMENTAL PROCEDURES

Reagents, strains, vectors

The *Escherichia coli* XL1-Blue (tet^r) and VCSM13 helper phage (kan^r) were purchased from Stratagene. All enzymes were purchased from Boehringer Mannheim and Promega. The phagemid vector pComb8 (Kang *et al.*, 1991) utilized to produce clones encoding TCRα chains fused to the cpVIII of M13 filamentous phage was generously provided by Denise Burton (The Scripps Research Institute). All molecular biology procedures were performed according to conventional techniques as described in *Molecular Cloning: a Laboratory Manual* (Sambrook *et al.*, 1989) unless indicated otherwise.

Peptides and antibodies

Poly 18 peptides were kindly provided by Dr Bhagirath Singh (University of Western Ontario, Canada). Bee venom phospholipase A_2 and bovine phospholipase A_2 were purchased from Sigma. Cell lines producing hamster anti-mouse TCRα (H28-710.16) and anti-TCRβ (H57-597 hamster IgG) were kindly provided by Dr Ralph Kubo (Cytel, La Jolla, CA, U.S.A.). Supernatants were concentrated by precipitation with 45% saturated ammonium chloride followed by dialysis, and the antibodies were purified by protein A chromatography.

The TCRα cDNA and clones

Full-length cDNA sequences encoding the A1.1 TCRα (Green *et al.*, 1991), 5C.C7 TCRα (Fink *et al.*, 1986) (provided by S. Hedrick, University of California, San Diego, CA, U.S.A.) and the 3B3 TCRα (Mori *et al.*, 1993) (T. Mikayama, unpublished sequence), respectively, were used as templates in the polymerase chain reaction (PCR) subcloning procedures. Fifty nanograms of cDNA template was mixed with 60 pmol of each primer (Table 1), 200 mM dNTP, Promega *Taq* polymerase buffer containing 1.5 mM $MgCl_2$ and five units of *Taq* polymerase in a final volume of 100 ml. Amplification was carried out for 20–30 cycles on a TwinBlock thermal cycler (Ericomp Inc., San Diego, CA, U.S.A.) under the following conditions: 1 min of denaturation at 94°C, 2 min of primer annealing at 50°C, 1 min of elongation at 72°C, followed by a final elongation at 72°C for 15 min. Amplified products were hydrolysed (4 hr at 37°C) at the *Xho*I and *Xba*I sites (25 U enzyme/μg PCR fragment) encoded by the primers, size-fractionated by agarose gel electrophoresis and purified using a Geneclean procedure (Bio 101 Inc., La Jolla, CA, U.S.A.). The PCR primers were designed to produce TCRα cDNA sequences that could be ligated into the pComb8 vector in frame with the pelB leader sequence at the 5′ end and fused to the N-terminus of the cpVIII protein on the 3′ end. The purified *Xho*I – *Xba*I insert was directionally cloned into the phagemid vector pComb8 at the *Xho*I and *Spe*I sites and transformants were screened with internal oligonucleotide probes in colony lifts. Plasmid DNA from positive clones were sequenced by the dideoxy method using Sequenase 2.0 (USB Corp., Cleveland, OH, U.S.A.), analysed using the MacVector analysis program (IBI, New Haven, CT, U.S.A.) confirming identity, orientation and in-frame cloning of TCRα cDNAs.

Recombinant phage preparation

Production of recombinant filamentous phage displaying TCRα chains was carried out essentially as previously described for the production of phage displaying immunoglobulin receptors (Kang *et al.*, 1991; Barbas and Lerner, 1991b). Briefly, *E. coli* XL1-Blue (Stratagene) cells were transformed with recombinant TCRα/pComb8

Table 1. The TCRα-specific primers

Variable region primers	A1
TCR Vα (sense) 5C.C7	5′-GAAGAGCTCGAGATGAAATCC TGAGT-3′
TCR Vα (sense) 3B3	5′-TGGACACTCGAGATGCAGAGGAACC
TCR Vα (sense)	5′-GCGCCGCTCGAGATGAAATCCTTGAGTGTTTTACTA GTGGTCCTGTGGCTCCAGTTAAACTGCGTGAGGAGC- CAGCAGCAAGTGCAGCAGAGTCCTGCA-3′
Constant region primers	
TCR-VJC	
TCR Cα (antisense)	5′-GCTGTCTCTAGAGCCACCGCCACCGTCGACG-TACACAGCAGGTTCTGGGTT-3
TCR-VJ	
Truncated Cα (antisense)	5′-CAGGAGTCTAGAGCCACCGCCACCGTCGACG-TTTGAAAGTTTAGGTTCATATC-3′

Restriction sites used

phagemid and selected on LB plates containing 100 μg/ml ampicillin. Fresh ampicillin-resistant colonies were grown in liquid culture on a shaker at 37°C in superbroth (SB) for 1–2 hr in the presence of ampicillin (50 μg/ml) to select for cells bearing phagemid and tetracycline (10 μg/ml) to induce the F′ episome. The cultures were then incubated with 1 mM isopropyl β-D-thiogalactoside (IPTG) for 1 hr to induce production of the TCR–cpVIII fusion protein and then superinfected with VCSM13 helper phage (10^{12} pfu) (Stratagene) to give a final phage/cell ratio of 10–20:1. The cells were shaken (37°C) for an additional 2 hr and then selected for helper phage induced antibiotic resistance (kanamycin, 70 μg/ml) and grown overnight, 37°C, 250 rpm. Phage supernatant was cleared by centrifugation of the cultures (4000 rpm in a GSA rotor, Sorvall, at 4°C). The phage were precipitated by adding 3% (w/v) NaCl and 4% (w/v) polyethylene glycol 8000 for 1 hr at 4°C and centrifuged at 9000 rpm in a Sorvall GSA rotor at 4°C. The phage pellets were resuspended in PBS to 1/25 of the original volume and aggregated phage and debris were removed by centrifugation for 5 min in a benchtop microcentrifuge. Phage supernatants were transferred to fresh tubes and stored in aliquots at −20°C.

Affinity selection panning

The panning procedure utilized to screen binding of TCR displayed on the surface of filamentous phage was a modification of the original protocol described by Parmley and Smith (1988). Wells of a microtitration ELISA plate (Immulon 2, Dynatech) were coated with peptide or phospholipase A_2 (Sigma) in sodium bicarbonate (0.1 M, pH 9.5) at 4°C overnight in a volume of 50 μl at the concentrations indicated in the figure legends. Antibodies were coated onto plates using Tris (50 mM, pH 9) at 4°C overnight in a volume of 50 μl at the concentrations indicated in the figure legends. The wells were washed twice with PBS (Dulbecco's, pH 7.4) and blocked by completely filling the well with 3% bovine serum albumin (BSA) (Sigma, fraction V) and incubated for 1 hr at 37°C. The blocking solution was flicked out, rinsed with PBS and 50 μl recombinant phage was added to each well (typically 10^{10}–10^{11} CFU unless otherwise indicated) and incubated for 2 hr at 37°C. The phage were then removed and the wells were washed 10× (pipetting up and down to wash) with TBS/Tween (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 0.05% Tween 20) over a period of 1 hr allowing the wells to remain filled between washings. Washing was done carefully to ensure similarity between samples, experiments and individuals. Adherent phages were eluted with 50 μl elution buffer (0.1 M HCl adjusted to Ph 2.2 with solid glycine, 1 mg/ml BSA) for 10 min at room temperature and neutralized with 6 ml of 1 M Tris base. Eluted phage were used to infect fresh *E. coli* XL1-Blue cells, OD_{600} = 1 (grown in the presence of 10 μg/ml tetracycline to induce F′ episome expression) for 15 min at 37°C followed by selection on LB/ampicillin plates to assess the number of bound recombinant phage per ml of eluted phage.

Immunoblot of TCRα–pVIII fusion proteins

Phagemid DNA of TCR-pComb8 or control constructs were transformed in XL1-Blue cells and transformants were inoculated in 25 ml 2×TY medium containing 50 μg/ml ampicillin. Protein expression was induced with 0.1 mM IPTG at an OD_{600} = 1 for 7 hr at 26°C. Cells were harvested by centrifugation, sonicated on ice for 1 min in PBS, 1 mM PMSF, 1% NP-40 and the resulting lysate was centrifuged (13,000g, 4°C) to remove insoluble debris. The soluble cellular proteins were analyzed by SDS–PAGE and immunoblotted with an anti-TCRα antibody (H28-710.16).

Presentation assay

The TCR specificity of A1.1 hybridoma cells for our poly 18-related peptide analogues was assessed by measuring the IL-2 responses of A1.1 cells to peptide-pulsed BALB/c (I-A^d) spleen cells (γ-irradiated 2000 rad). Supernatants were collected following overnight peptide stimulation and the relative production of IL-2 was assessed in a CTLL assay. Starved CTLL cells were washed and

incubated in various dilutions of sample or control supernatants (100 μl) in 96-well plates (Costar) at 37°C for 16 hr. The cells were then pulsed with [^{3}H]-thymidine (1 μCi/well) and harvested an additional 4–6 hr later. The IL-2-dependent proliferation was assessed by the relative level of [^{3}H]-thymidine incorporation (cpm).

Hybridization of replica plates

Following two rounds of affinity selection panning (described above), ampicillin-resistant CFUs were blotted onto replica filters as described (Sambrook *et al.*, 1989). The replica filters were denatured (0.5 N NaOH, 1.5 M NaCl) and neutralized [1.5 M NaCl, 0.5 M Tris–HCL (pH 7.4)] and then fixed (80°C, 2 hr). The filters were then hybridized with a 5′ ^{32}P-labelled A1.1 TCR Vα specific probe (5′-GAAGAGCTCGAGATGAAATC-CTTGAGT-3′).

RESULTS

The TCR phage display

We previously presented evidence suggesting that the TCRα molecule from the A1.1 T cell hybridoma (Vα1.2, JαTA65) may bind directly to antigenic peptides for which this T cell is specific (Bissonnette *et al.*, 1991; Green *et al.*, 1991). The M13 filamentous phage display system (Kang *et al.*, 1991; Barbas *et al.*, 1991a; Barbas and Lerner, 1991b) provided the means to directly assess the ability of A1.1 TCRα polypeptides to specifically bind to peptide antigens. We utilized the phagemid vector, pComb8 (Kang *et al.*, 1991) to generate recombinant phage-displaying TCR proteins fused to the N-terminus of the coat protein VIII (cpVIII). The cpVIII molecules form the capsid coat during phage assembly (utilizing about 2500 cpVIII proteins per phage particle) allowing for multiple receptors to be displayed on the surface of the recombinant phage (Felici *et al.*, 1991; Greenwood *et al.*, 1991; Kang *et al.*, 1991). This multivalent expression can greatly enhance the avidity, which facilitates the study of binding interactions with moderate affinity.

The PCR-generated fragments encoding the VJ- or VJC-domains of the A1.1 TCRα molecule were subcloned into the pComb8 phagemid vector to generate recombinant phage displaying multiple copies of the A1.1 TCRα molecule on the surface. This is illustrated in Fig. 1 (see Experimental procedures). Extracts of cells transformed with various recombinant or control phagemids were subjected to SDS–PAGE and immunoblotted with an anti-TCR Cα specific antibody (H28.710.16) to confirm that TCRα/cpVIII fusion proteins were produced. As shown in Fig. 2A, a protein that migrated with the predicted molecular weight for the TCR–VJC/cpVIII fusion protein, Mr 36.5 kDa, was detected from extracts of cells transformed with the A1.1 TCR–VJCα construct. A second protein of M, 34.5 kDa was also detected, which probably represents a partially degraded form of the fusion protein, since it was not observed in the other lanes. No bands were observed from extracts of cells transformed with either the parental phagemid (pComb8, vector only) or the A1.1 TCR VJα construct, which both lack the Cα domain recognized by the H28.710.16 antibody used for detection.

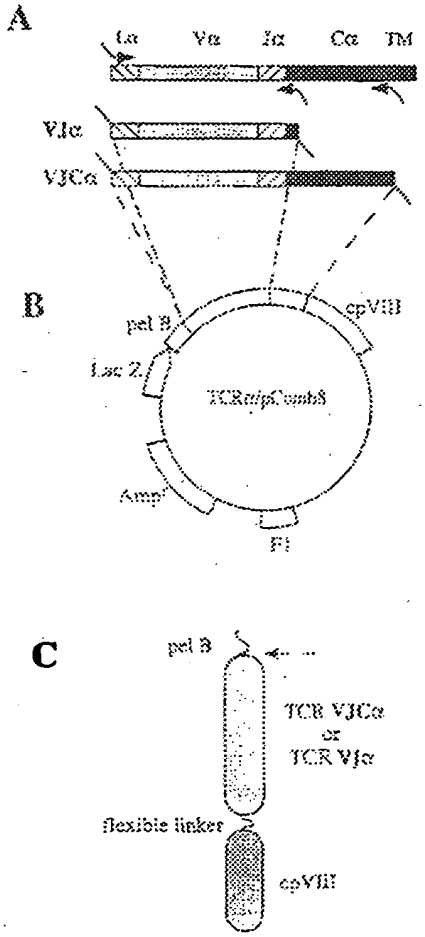


Fig. 1. Schematic protocol for the generation of chimeric constructs. (A) Full-length cDNA sequences encoding various TCRα chains were used as templates in the polymerase chain reaction (PCR) subcloning procedures. The 5′ PCR primers (arrows) contained an *Xho*I site designed to generate a PCR fragment that could be ligated into the pComb8 vector in frame with the pelB leader sequence. The 3′ primers (arrows) contained a sequence encoding a flexible linker (GGGS) and an *Xba*I site to generate a fusion gene with the major cpVIII upon ligation. The PCR fragments were digested with *Xba*I and *Xho*I, purified and directionally cloned into the phagemid vector pComb8. (B) The recombinant TCRα–pComb8 phagemids contained an amp′ gene utilized to facilitate drug selection of recombinant phage. Production of the recombinant TCRα/cpVIII fusion protein was driven by the inducible LacZ promoter following transformation of competent XL1-Blue cells (Stratagene). Transport into the periplasmic space during production was mediated by the pelB leader sequence. (C) The predicted structure of the recombinant TCRα/cpVIII fusion proteins consisted of a TCRα polypeptide and a flexible linker (GGGS) fused to the N-terminus of the cpVIII molecule. The pelB leader sequence at the N-terminal end is cleaved off following transport into the periplasmic space. The recombinant fusion protein was incorporated into the viral capsid coat during phage assembly to generate recombinant M13 phage displaying TCR molecules on the surface.

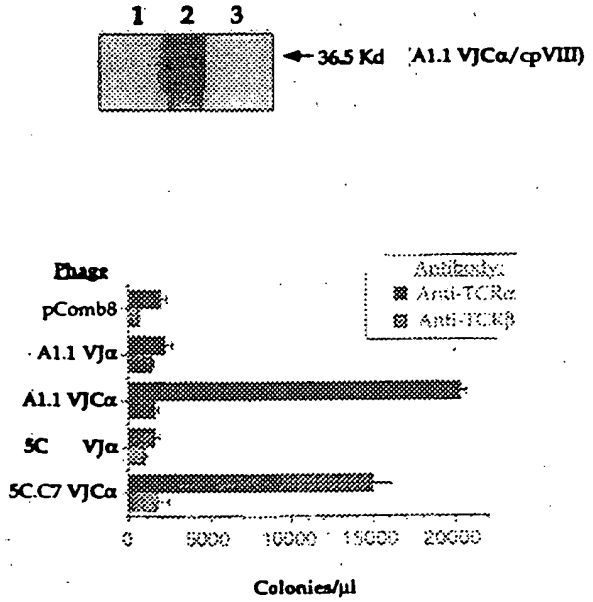


Fig. 2. Expression of TCRα/pComb8 recombinant fusion proteins. (A) Western blot analysis to detect expression of A1.1-TCR–VJCα/cpVIII fusion proteins in *E. coli*. Lysates from recombinant TCRα–pComb8 (lane 1: A1.1-TCR VJ; lane 2: A1.1-TCR–VJC) or parental pComb8 (lane 3: vector only) phagemid transformed XL1-Blue cells were subjected to SDS–PAGE and immunoblotted with an anti-TCR Cα antibody (H28.710.16). The position of the band migrating at the predicted molecular weight of the TCR–VJCα/cpVIII fusion protein (M, 36.5 kD) is indicated by the arrow. (B) Display of TCR–VJCα polypeptides on the surface of recombinant phage. Binding of recombinant phage to anti-TCR antibody was utilized to detect the presence of TCR on the phage surface. Recombinant TCRα–pComb8 phage or parental pComb8 phage (indicated in the figure) were panned on plates coated with antibodies specific for TCR Cα (H28.710.16, solid bars) or TCR Vβ6 (H57-597, hatched bars). The number of bound phage was assessed by counting the number of eluted ampicillin resistant CFU following extensive washing as described in the experimental procedures.

The presence of TCRα molecules displayed on the surface of recombinant phage was assessed using anti-TCR antibodies. Recombinant phage were "panned" on plastic dishes (Parmley and Smith, 1988) coated with either anti-TCR Cα (H28-710.16) or anti-TCR Cβ (H57-597) antibodies (Fig. 2B). Following extensive washing, the bound phage were eluted with a low pH buffer. Subsequent infection of host cells with eluted phage was readily achievable as the cpIII proteins, utilized for adsorption to the F′ episome and infection of *E. coli*, remain intact following acid elution. Since the pComb8 phagemid vector contains an *amp*^{r} gene, the number of bound phage could be assessed by counting the number of ampicillin-resistant CFU following infection of host cells with eluted phage. In addition to recombinant phage displaying A1.1 TCRα chains, we also constructed phage displaying the TCRα chains from the 5C.C7 T cell hybridoma (specific for pigeon cytochrome C peptide restricted to I-E^k) (Fink *et al.*, 1986). As shown in Fig. 2B, both the A1.1 TCR–VJCα as well as the 5C.C7-TCR–VJCα recombinant phage were readily bound by the anti-TCRα but not the anti-TCRβ antibodies. As predicted, the parental pComb8 phage (vector only) and recombinant TCR–VJα phage showed no binding to either antibody as these phage lack the TCR Cα determinant recognized by H28.710.16 antibody. Thus, polypeptides bearing TCR Cα specific determinants were effectively displayed on the surface of recombinant phage particles, indicating that TCRα/cpVIII fusion proteins could be effectively incorporated into the capsid coat during phage assembly.

Specificity of TCR displayed on recombinant phage

The A1.1 T hybridoma cells recognize a synthetic polypeptide with the sequence poly [EYK(EYA)$_5$] (poly 18) presented by I-A^d (Fotedar *et al.*, 1985). The antigenic fine specificity of A1.1 cells for a series of poly 18-related peptide analogues was assessed by measuring the IL-2 responses to peptide pulsed BALB/c (I-A^d) spleen cells. As shown in Fig. 3A, A1.1 cells produced cytokine in response to two of the peptides, EYK(EYA)$_4$ and EYK-(EYA)$_4$EYK. Peptide analogues, substituted with alanines at residues 3 or 10, failed to stimulate A1.1 cells. These results are in agreement with those reported by others for this cell line (Fotedar *et al.*, 1985). The antigenic fine specificity of the A1.1 hybridoma cells for the four peptides described above provided a model with which to compare specificity of TCRα molecules displayed on recombinant phage.

The poly 18-related peptide analogues utilized to characterize the specificity of A1.1 hybridomas were coated onto 96-well plates for panning experiments to assess the binding capacity and specificity of the TCRα molecules displayed on the phage surfaces. Bound phage were extensively washed, eluted and then quantitated by counting the number of ampicillin-resistant CFUs following infection of host cells. Representative of a series of 10 binding experiments, Fig. 3B shows that recombinant phage expressing A1.1 TCR–VJCα preferentially bound to the antigenic peptides, EYKEYAEYAEYAEYAEYK and EYKEYAEYAEYAEYA, but not to non-antigenic peptides with an alanine substitution at residues 3 (EY*A*-EYAEYAEYAEYA) or 10 (EYKEYAEYA*A*YAEYA). The parental pComb8 phage (vector only) and recombinant phage displaying 5C.C7-TCR–VJCα showed no specific binding to any of the four poly 18 peptide analogues (Figs 3B and 4A). Recombinant phages displaying A1.1 TCR VJα, which lack the Cα-domain, also showed preferential binding to the antigenic peptides, indicating that the specific binding was mediated by the TCR Vα- and Jα-region domains (Fig. 4B). The observed preferential binding to antigenic peptide was reproducible and consistent over a wide range of phage titres when we collected the quantitative data from seven different experiments utilizing several different recombinant phage preparations from two different recombinant constructs (Fig. 4C). Thus, direct binding by phage displaying A1.1 TCRα to peptides reflects the antigenic specificity of the

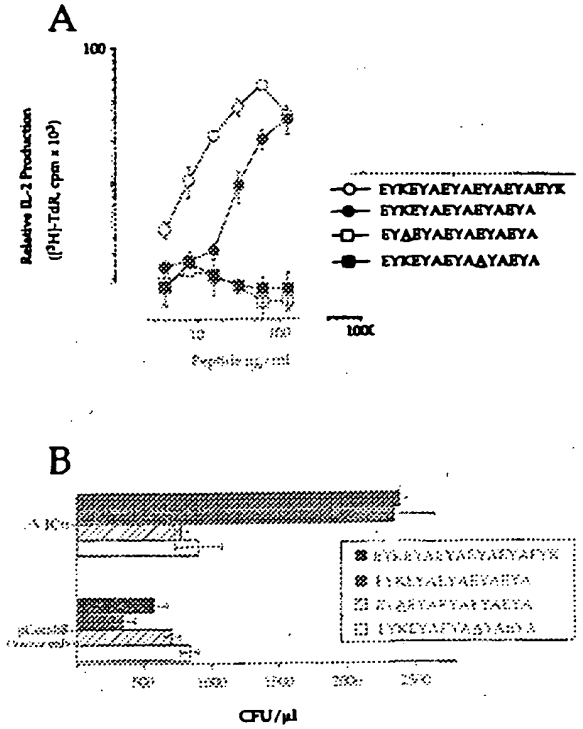


Fig. 3. Antigenic fine specificity of A1.1 hybridoma cells and A1.1-TCRα Recombinant phage. (A) Fine specificity of A1.1 hybridoma cells. The ability of poly 18 related peptide analogues to induce an IL-2 response was utilized to assess the fine specificity of A1.1 hybridoma cells. The BALB/c spleen cells (I-A^d) were pulsed with the indicated poly 18 related peptide analogues and co-cultured with A1.1 hybridoma cells as described in the experimental procedures. The relative amount of IL-2 produced in response to the peptide antigens is indicated in the *Y*-axis as the amount of tritiated thymidine uptake in a CTLL assay. (B) Comparative fine specificity of peptide binding by recombinant A1.1-TCR phage. The binding capacity of recombinant phage displaying A1.1-TCR–VJCα was assessed using the same four poly 18 related peptide analogues used to assess the fine specificity of the A1.1 hybridoma cells. Plates were coated with the indicated peptide analogues followed by panning with recombinant phage displaying A1.1 TCR–VJCα. The number of bound phage following extensive washing was assessed by determining the number of eluted ampicillin resistant CFU.

A1.1 hybridoma from which the TCRα cDNA sequence was derived (see Fig. 3A), despite the absence of class II MHC presentation.

Affinity selection of specific TCRα

To assess the capacity for affinity selection of phage displaying specific TCR polypeptides, recombinant phage displaying A1.1 TCR–VJCα were mixed 1:100 or 1:1000 with phage displaying 5C.C7-TCR–VJCα and then subjected to two rounds of affinity enrichment on plates coated with antigenic poly 18 peptide analogues. The eluted phage were used to infect host cells and colonies from ampicillin plates were lifted onto nitrocellulose filters. The colony-lifts were lysed and hybridized with a DNA probe specific for the A1.1 TCRα sequence. We found that 85% (468/554) of the colonies from the 1:100 mix and 76% (113/149) of the colonies from the 1:1000 mix contained A1.1 TCRα specific sequences (Fig. 5). Thus, two rounds of selection enriched the phage displaying A1.1 TCRα by over 500- and 3000-fold, respectively. These results suggest that the phage-display system can be utilized to select for T cell receptors capable of binding directly to peptide ligands and has potential to be utilized as a gene identifying system for TCRα chains with relatively high affinity for their peptide ligand.

The TCRα binding directly to protein antigen

While phage displaying 5C.C7-TCR–VJCα failed to bind to the poly 18 peptide analogues bound by A1.1 TCRα recombinant phage, these phage also failed to bind antigenic cytochrome C peptide, in repeated attempts (data not shown). Thus, differences in detectable binding of TCR Vα to specific antigenic peptides may represent variance in the particular TCR–ligand interactions such that only a subset of TCR, including that on A1.1 cells, have direct interactions strong enough to be detectable in this system. Accordingly, we constructed recombinant phage displaying TCRα from the 3B3 T cell hybridoma, which has previously been suggested to express a TCRα chain product capable of direct interaction with antigen (Mori *et al.*, 1993). The 3B3 hybridoma was generated from BALB/c mice immunized with bee venom phospholipase A_2 (PLA_2) and is specific for a bee venom PLA_2 derived peptide (residues 19–34) in the context of I-A^d (Mori *et al.*, 1993). Two aspects of 3B3 TCR binding characteristics made it particularly interesting for phage display studies of direct antigen-binding TCRα chains. Firstly, three-dimensional structural analysis of bee venom PLA_2 indicated that the antigenic peptide residues associated with 3B3 TCR binding are exposed to the outer surface of the PLA_2 molecule in the form of a flexible loop (Scott *et al.*, 1990). Secondly, previous peptide competition studies of 3B3 TCR specificity suggested that the TCRα chain may be able to bind directly to unprocessed bee venom PLA_2 (Mori *et al.*, 1993). Therefore, we constructed recombinant phage displaying 3B3-TCR–VJCα to test the ability of this TCRα chain to directly bind to unprocessed bee venom PLA_2 protein in a specific manner.

The PCR-generated fragments encoding the 3B3-TCR–VJCα polypeptide were subcloned into the pComb8 vector for production of recombinant phage. Bee venom PLA_2 was coated onto 96-well plates for panning experiments to assess the binding capacity and specificity of the 3B3 TCRα chains displayed on the phage surface. We observed binding of recombinant 3B3-TCR–VJCα phage, but not parental pComb8 phage (vector only), to bee venom PLA_2 coated wells (Fig. 6A). Furthermore, binding of recombinant 3B3-TCR–VJCα phage to BSA (3% BSA, used for blocking) was not observed. Moreover, the recombinant phage displaying 3B3 TCR–VJCα did not bind to bovine PLA_2 (Fig. 6B), which lacks the exposed hydrophilic loop containing the

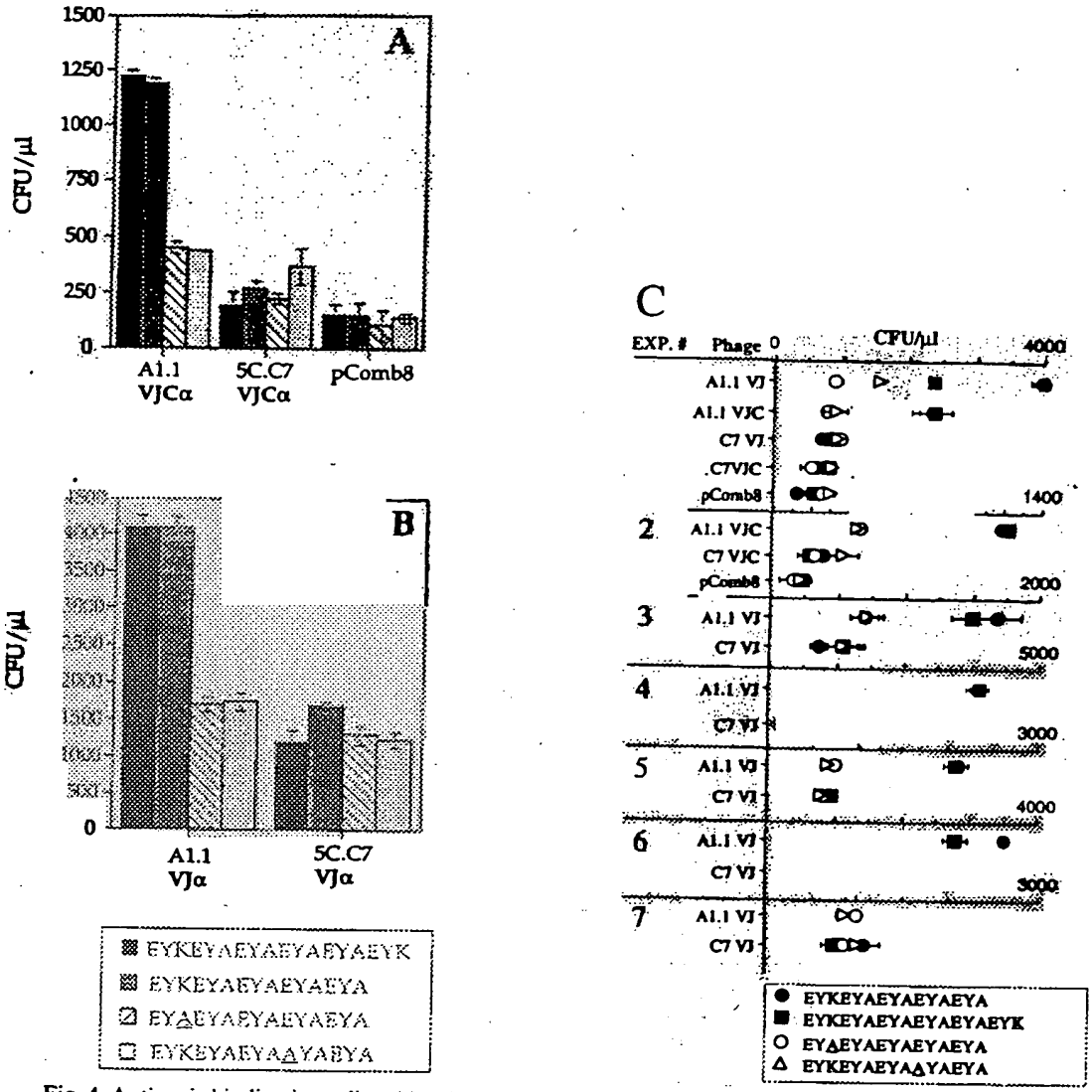


Fig. 4. Antigenic binding is mediated by the VJ regions of specific TCR-recombinant phage. (A) The binding capacity of parental pComb8 phage or recombinant phage displaying A1.1-TCR–VJCα or 5C.C7-TCR–VJCα for the indicated poly 18 related peptide analogues was assessed using the panning protocol described in the experimental procedures. Plates were coated with the indicated peptide analogues and panned with the indicated phage. The number of bound phage following extensive washing was assessed by determining the number of ampicillin resistant CFU/μl. (B) The binding capacity of recombinant phage displaying TCR VJα (truncated C-region) from the A1.1 or 5C.C7 hybridoma cells to the four indicated poly 18 related peptide analogues was determined as described above to assess whether the VJ-regions were sufficient to mediate the specific binding to antigenic peptides. (C) The relative binding specificity of several recombinant phage preparations, displaying TCR–VJCα or TCR VJα, for the indicated poly 18 related peptide analogues are compared from seven different panning experiments as described above. The experiments described in (A) and (B) are included for comparison.

antigenic peptide sequence associated with 3B3 TCR binding (Dijkstra *et al.*, 1981; Mori *et al.*, 1993). The observed preferential binding to antigenic bee venom PLA_2 was reproducible and consistent over a wide range of phage titres when we collected the quantitative data from five different experiments utilizing several different recombinant phage preparations (Fig. 6C). Thus, the phage display system enabled us to assess the novel capacity of a TCRα chain to bind directly to antigenic protein in the absence of MHC.

In order to further assess the specificity of the TCRα recombinant phages for their respective antigenic ligands, we did parallel panning experiments with the A1.1 TCR–VJCα and 3B3 TCR–VJCα recombinant phage (Fig. 7). Phage displaying A1.1 TCR–VJCα showed binding to the antigenic poly 18 peptide analogue (EYK-EYAEYAEYAEYAEYK) but not to bee venom PLA_2. In contrast, phage displaying 3B3 TCR–VJCα demonstrated specific binding to bee venom PLA_2 but not to the poly 18 peptide analogue. Thus, the TCR phage-

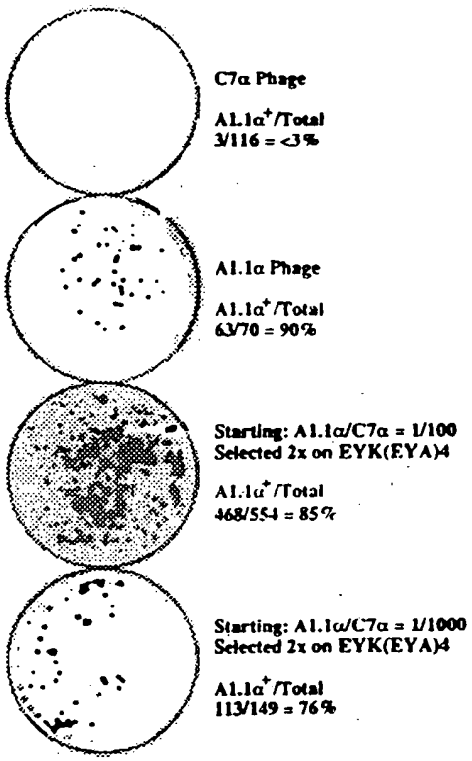


Fig. 5. Affinity selection of specific TCR on recombinant phage. A mixing experiment was done to assess the capacity for affinity selection of peptide specific TCR displayed on the surface of recombinant A1.1-TCRα–pComb8 phage. The A1.1-TCR–VJCα phages were mixed 1:100 or 1:1000 with 5C.C7-TCR–VJCα phage and subjected to two rounds of affinity selection by panning as described in the experimental procedures. Colony lifts were hybridized with a radiolabeled A1.1-TCRα specific oligonucleotide probe. Autoradiographs of the hybridized replica filters are shown in panels A–D. Autoradiographs of control colony lifts from plates inoculated with unmixed 5C.C7 TCRα recombinant phage (panel A) or unmixed A1.1-TCRα recombinant phage (panel B) are shown for comparison with plates from mixing experiments starting with A1.1α/5C.C7α ratios of 1:100 (panel C) or 1:1000 (panel D) prior to the two rounds of affinity selection.

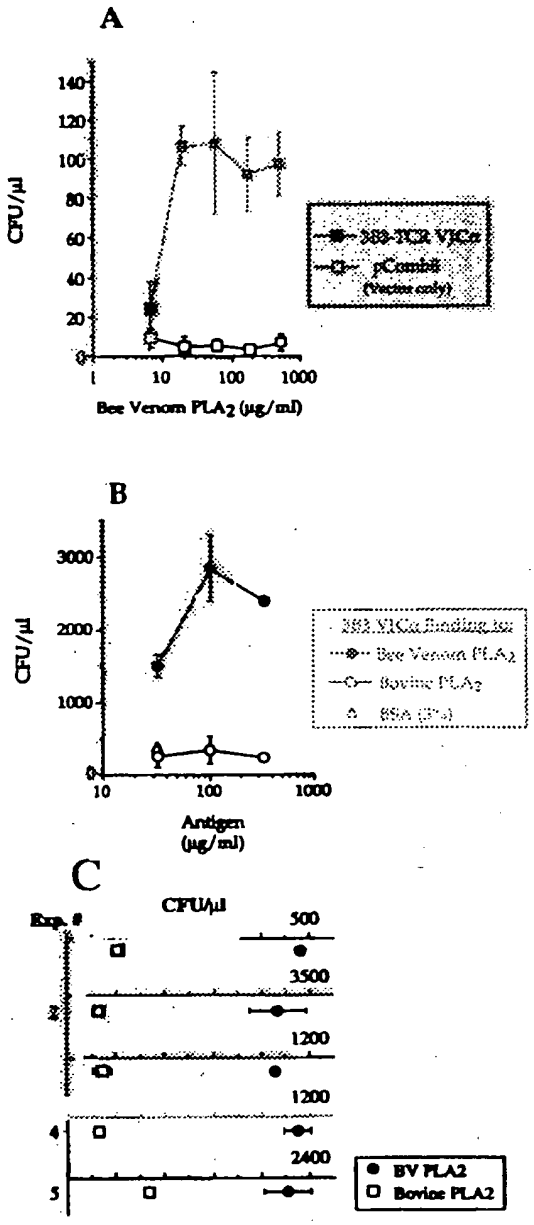
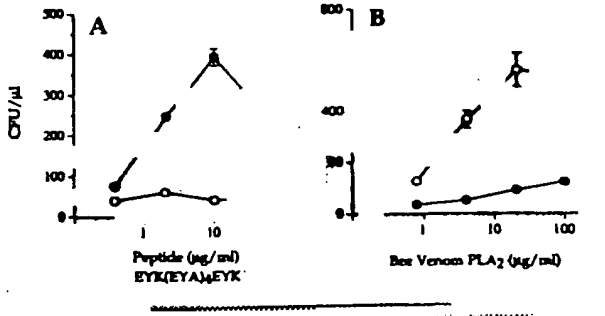


Fig. 6. Antigenic specificity of recombinant phage displaying TCRα polypeptide from 3B3 hybridoma cells. (A) The binding capacity of parental pComb8 phage or recombinant phage displaying TCR–VJCα from the 3B3 hybridoma cell to bee venom PLA_2 was assessed using the panning protocol described in the experimental procedures. Plates were coated with the indicated concentration of bee venom PLA_2 and then panned with 10^9 CFU of recombinant phage displaying 3B3-TCR–VJCα or parental pComb8 phage. The number of bound phage following extensive washing was assessed by determining the number of ampicillin resistant CFU/μl. (B) The specificity of recombinant phage displaying TCR–VJCα from the 3B3 hybridoma cell for bee venom PLA_2 and bovine PLA_2 was assessed using the panning protocol. The number of bound phage following panning with 10^{10} CFU of recombinant phage and extensive washing was assessed by determining the number of ampicillin resistant CFU/μl. (C) The relative binding specificity of several recombinant phage preparations, displaying 3B3-TCR–VJCα for bee venom PLA_2 or bovine PLA_2 are compared from five different panning experiments as described above.

display system described here facilitated the study of TCRα chains capable of binding directly to antigenic epitopes in a specific manner.

DISCUSSION

The phage-display system (Smith, 1985) has proved to be extremely useful for the study of specific receptor–ligand interactions (Winter *et al.*, 1994). The advantage of this system for immunological studies is that the display of receptors on recombinant phage provides a means to link directly antigen recognition structures and the genetic instructions encoding the receptor (Kang *et al.*, 1991). Consequently, the phage display system has been invaluable for the characterization and selection of antibody specificities for antigenic ligands. Here, we have shown that this system can be utilized further to study specific binding interactions of constituent TCR poly-



peptides to antigenic ligands. Specifically, the phage-display system facilitated a novel demonstration of direct binding by TCRα chains to peptide or protein antigens in a specific manner. These results indicate that for some TCRα chains the binding affinity to antigenic peptide may be sufficiently high that it can be detected in the absence of the TCRβ chain, and in an MHC-independent manner. It has been postulated previously that TCR binding to peptide by Vα- and Vβ-subunits may be relatively autonomous and that the resulting TCR–peptide interactions can dramatically influence TCR–MHC interactions; indicating the "primacy" of TCR–peptide interactions (Ehrich *et al.*, 1993). Our results extend these findings by demonstrating that the dominant interactions of certain TCRα chains for peptide antigens may be sufficiently high that they can be analysed independently. However, these interactions are quite unusual in that they do not require the expression of the second TCR subunit or normal MHC and coreceptor interactions. These results raise the concern that this model does not reflect typical TCR–ligand interactions. Indeed, gene transfer studies of A1.1 TCRα chain sequences indicated that the expression of the α chain alone was not sufficient to confer responsiveness to peptide antigen alone or in the context of I-A^d (H. Zheng, A. Fotedar and D.R. Green; unpublished data). Furthermore, activation of A1.1 hybridoma cells required specific peptide antigen to be presented in the context of the appropriate MHC (I-A^d) (Fotedar *et al.*, 1985). The requirement for MHC presentation and natural TCRαβ conformation for functional interaction of cell surface expressed receptors may be essential for bringing the respective binding sites into close enough proximity to establish the specific binding interaction between the residues of opposite charge which mediate binding. The specificity data presented here (Fig. 3 and Fig. 4 and Fig. 7) suggest that these requirements may not be necessary for some receptors to detect specific binding in non-cellular TCR phage-display systems. Therefore, the data presented here do not suggest that the TCRα chain alone would be sufficient for T cell recognition and activation, but rather that the TCR phage-display system provides a means to explore the specific binding interactions that mediate TCR binding to antigen at a constituent level. While this study does not refute commonly held theories of TCR–antigen interactions, this system does provide a means to explore the innate specific binding interactions contributed by the constituent subunits of TCR complexes for the peptide component of MHC–peptide ligands. In the absence of detailed structural data for the TCR/peptide/MHC trimolecular complex, models such as this can provide important insight into the nature of the elemental interactions that constitute TCR binding to MHC–peptide complexes. Furthermore, such information should facilitate interpretation of the structural data for the trimolecular complexes when this becomes available.

Previous studies indicate that TCR–peptide binding interactions are very specific, since single amino acid changes had profound effects on binding (Engel and Hedrick, 1988; Danska *et al.*, 1990). In accordance with these results, the direct binding of phage-displayed TCRα chains to antigen appeared to be specific. Recombinant phage-displaying A1.1 TCRα chains bound to antigenic peptide analogues but not to mutated peptide analogues that failed to activate A1.1 hybridoma cells (Fig. 3 and Fig. 4). Additionally, the cross-specificity experiment indicated that recombinant phage-displaying A1.1 TCR–VJCα or 3B3-TCR–VJCα bound their respective antigens but did not cross-react (Fig. 7).

The failure of repeated attempts to demonstrate specific binding of recombinant phage-displaying 5C.C7 TCRα to cytochrome C peptide indicated that only a subset of TCR Vα have the capacity for direct interactions with antigen strong enough to be detectable in this system. Previous studies of 5C.C7 TCR interactions with peptide antigen indicate that specific binding was mediated by CDR3 residues from both the α and β chains (Jorgensen *et al.*, 1992a, 1992b). Thus, binding to antigen by receptors such as 5C.C7 TCR may require the combined binding interactions of the α and β chains to mediate strong specific binding, while antigen binding by receptors such as A1.1 and 3B3 TCR may be dominated by specificities mediated predominantly by the TCRα chain.

The phage-display system described here extends the study of T cell recognition by providing a means to simplify the study of the specific interactions that constitute the overall receptor–ligand interactions. The power of the phage-display system is that recombinant phage can be amplified and subjected to multiple rounds of affinity selection as they contain the genetic instructions for production of the receptor molecule. A system for selection and cloning of antigen-binding TCR molecules will be useful in exploring TCR binding to various antigens and should provide a means to more closely examine TCR–ligand interactions.

Acknowledgements—We thank Dr K. Ishizaka and Dr A. Kang for helpful discussions. We thank Dr N. Gascoigne for critically reviewing the manuscript. We also thank Dr D. Burton (The Scripps Research Foundation) for the pComb8 phage display expression vector and Dr S. Hedrick (University of California, San Diego, U.S.A.) for the 5C.C7 TCRα cDNA sequence. This work was supported in part by National Institutes of Health grant GM52735. D.M.L. was a Myasthenia Gravis Foundation Fellow.

REFERENCES

Ajitkumar P., Geier S. S., Kesari K. V., Borriello F., Nakagawa M., Bluestone J., Saper M. A., Wiley D. C. and Nathenson S. G. (1988) Evidence that multiple residues on both the α helices of the class I MHC molecule are simultaneously recognized by the T cell receptor. *Cell* **54**, 47–56.

Barbas C. F., Kang A. S., Lerner R. A. and Benkovic S. J. (1991a) Assembly of combinatorial antibody libraries on phage surfaces: The gene III site. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7978–7982.

Barbas C. F. and Lerner R. A. (1991b) Combinatorial immunoglobulin libraries on the surface of phage (Phabs): Rapid selection of antigen-specific Fabs. *Meth. Compan. Meth. Enzymol.* **2**, 119–124.

Bissonnette R., Zheng H., Kubo R. T., Singh B. and Green D. R. (1991) A T helper cell hybridoma produces an antigen-specific regulatory activity. *J. Immun.* **146**, 2898–2907.

Chien Y.-H. and Davis M. M. (1993) How $\alpha\beta$ T-cell receptors 'see' peptide/MHC complexes. *Immun. Today* **14**, 597–602.

Danska J. S., Livingstone A. M., Paragas V., Ishihara T. and Fathman C. G. (1990) The presumptive CDR3 regions of both T cell receptor α and β chains determine T cell specificity for myoglobin peptides. *J. exp. Med.* **172**, 27–33.

Davis M. and Chien Y.-H. (1993) Topology and affinity of T-cell receptor mediated recognition of peptide–MHC complexes. *Curr. Opin. Immun.* **5**, 45–49.

Davis M. M. and Bjorkman P. J. (1988) T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402.

Dijkstra B. W., Kalk K. H., Hol W. G. J. and Drenth J. (1981) Structure of bovine pancreatic phospholipase A_2 at 1.7 Å resolution. *J. molec. Biol.* **147**, 97–123.

Ehrich E. W., Devaux B., Rock E. P., Jorgensen J. L., Davis M. M. and Chien Y. H. (1993) T cell receptor interaction with peptide/major histocompatibility complex (MHC) and superantigen/MHC ligands is dominated by antigen. *J. exp. Med.* **178**, 713–722.

Engel I. and Hedrick S. M. (1988) Site-directed mutations in the VDJ junctional region of a T cell receptor β chain cause changes in antigenic peptide recognition. *Cell* **54**, 473–484.

Esch T. R. and Thomas D. W. (1990) Ia-independent binding of peptide antigen to the T cell receptor. *J. Immun.* **145**, 3972–3978.

Felici F., Castagnoli L., Musacchio A., Jappelli R. and Cesareni G. (1991) Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J. molec. Biol.* **222**, 301–310.

Fink P. J., Matis L. A., McKlligott D. L., Bookman M. and Hedrick S. M. (1986) Correlations between T-cell specificity and the structure of the antigen receptor. *Nature* **321**, 219–226.

Fotedar A., Boyer M., Smart W., Widtman J., Fraga E. and Singh B. (1985) Fine specificity of antigen recognition by T cell hybridoma clones specific for poly-18: a synthetic polypeptide antigen of defined sequence and conformation. *J. Immun.* **135**, 3028–3033.

Gascoigne N. R. and Ames K. T. (1991) Direct binding of secreted T-cell receptor beta chain to superantigen associated with class II major histocompatibility complex protein. *Proc. natn. Acad. Sci. U.S.A.* **88**, 613–616.

Green D. R., Bissonnette R., Zheng H., Onda T., Echeverri F., Mogil R. J., Steele J. K., Voralia M. and Fotedar A. (1991) Immunoregulatory activity of the T-cell receptor α chain demonstrated by retroviral gene transfer. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8475–8479.

Greenwood J., Willis A. E. and Perham R. N. (1991) Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from *Plasmodium falciparum* circumsporozoite protein as antigens. *J. molec. Biol.* **220**, 821–827.

Haurum J. S., Arsequell G., Lellouch A. C., Wong S. Y. C., Dwek R. A., McMichael A. J. and Elliot T. (1994) Recognition of carbohydrate by major histocompatibility complex class I-restricted glycopeptide-specific cytotoxic T lymphocytes. *J. exp. Med.* **180**, 739–744.

Jorgensen J. L., Esser U., Groth B. F. d. S., Reay P. A. and Davis M. M. (1992a) Mapping T-cell receptor–peptide contacts by variant peptide immunization of single-chain transgenics. *Nature* **355**, 224–230.

Jorgensen J. L., Reay P. A., Ehrich E. W. and Davis M. M. (1992b) Molecular components of T-cell recognition. *A. Rev. Immun.* **10**, 835–873.

Kang A. S., Barbas C. F., Janda K. D., Benkovic S. J. and Lerner R. A. (1991) Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4363–4366.

Kuchroo V. K., Byrne M. C., Atsumi Y., Greenfield E., Connolly J. B., Whitters M. J. Jr, R. M. O. H. Collins M. and Dorf M. E. (1991) T-cell receptor α chain plays a critical role in antigen-specific suppressor cell function. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8700–8704.

Michaëlsson E., Malmström V., Reis S., Engström Å., Burkhardt H. and Holmdahl R. (1994) T cell recognition of carbohydrates on type II collagen. *J. exp. Med.* **180**, 745–749.

Mori A., Thomas P., Tagaya Y., Iijima H., Grey H. and Ishizaka K. (1993) Epitope specificity of bee venom phospholipase A_2-specific suppressor T cells which produce antigen-binding glycosylation inhibiting factor. *Int. Immun.* **5**, 833–842.

Parmley S. F. and Smith G. P. (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**, 305–318.

Peccoud J., Dellabona P., Allen P., Benoist C. and Mathis D. (1990) Delineation of antigen contact residues on an MHC class II molecule. *EMBO J.* **13**, 4215–4223.

Porcelli S., Morita C. T. and Brenner M. B. (1992) CD1β restricts the response of human CD4$^-$8$^-$ T lymphocytes to a microbial antigen. *Nature* **360**, 593–597.

Rao A., Faas S. J. and Cantor H. (1984a) Activation specificity of arsonate-reactive T cell clones. *J. Exp. Med.* **159**, 479–494.

Rao A., Ko W. W.-P., Faas S. J. and Cantor H. (1984b) Binding of antigen in the absence of histocompatibility proteins by arsonate-reactive T-cell clones. *Cell* **36**, 879–888.

Sambrook J., Fritsch E. F. and Maniatis T. (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.

Scott D. L., Otwinowski Z., Gelb M. H. and Sigler P. B. (1990) Crystal structure of bee-venom phospholipase A_2 in a complex with a transition-state analogue. *Science* **250**, 1563–1566.

Siliciano R. F., Hemesath T. J., Pratt J. C., Dintzis R. Z., Dintzes H. M., Acuto O., Shin H. S. and Reinherz E. L. (1986) Direct evidence for the existence of nominal antigen binding sites on T cell surface Ti α–β heterodimers of MHC-restricted T cell clones. *Cell* 47, 161–171.

Smith G. P. (1985) Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science* 228, 1315–1317.

White J., Herman A., Pullen A. M., Kubo R., Kappler J. W. and Marrack P. (1989) The V beta-specific superantigen staphylococcal enterotoxin B: stimulation of mature T cells and clonal deletion in neonatal mice. *Cell* 56, 27–35.

White J., Pullen A., Choi K., Marrack P. and Kappler J. W. (1993) Antigen recognition properties of mutant Vβ3$^+$ T cell receptors are consistent with an immunoglobulin-like structure for the receptor. *J. exp. Med.* 177, 119–125.

Winter G., Griffiths A. D., Hawkins R. E. and Hoogenboom H. R. (1994) Making antibodies by phage display technology. *A. Rev. Immun.* 12, 433–455.

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 12654–12658, December 1994
Immunology

Functional three-domain single-chain T-cell receptors

SHAN CHUNG†‡, KAI W. WUCHERPFENNIG†, STEVEN M. FRIEDMAN§, DAVID A. HAFLER¶, AND JACK L. STROMINGER†

†Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138; §Department of Medicine, Cornell University Medical College, New York, NY 10021; and ¶Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

Contributed by Jack L. Strominger, July 19, 1994

ABSTRACT **T-cell receptors (TCRs) are membrane anchored heterodimers structurally related to antibody molecules. Single-chain antibodies can be engineered by linking the two variable domains, which fold properly by themselves. However, proper assembly of the variable domains of a human TCR (V_α and V_β) that recognize the HLA-DR2b/myelin basic protein-(85–99) peptide complex was critically dependent on the addition of a third domain, the constant region of the TCR β chain (C_β), to the single-chain construct. Single-chain molecules with the three-domain design, but not those with the two-domain design, expressed in a eukaryotic cell as chimeric molecules linked either to glycosyl phosphatidylinositol or to the transmembrane/cytoplasmic domains of the CD3 ζ chain were recognized by a conformation-sensitive monoclonal antibody. The chimeric three-domain single-chain TCR linked to CD3 ζ chain signaled in response to both the specific HLA-DR/peptide and the HLA-DR/superantigen staphylococcal enterotoxin B complexes. Thus, by using this three-domain design, functional single-chain TCR molecules were expressed with high efficiency. The lipid-linked single-chain TCR was solubilized by enzymatic cleavage and purified by affinity chromatography. The apparent requirement of the constant domain for cooperative folding of the two TCR variable domains may reflect significant structural differences between TCR and antibody molecules.**

T-cell receptor (TCR) recognition of antigen fragments presented by major histocompatibility complex (MHC) molecules is a critical step in the initiation of a specific immune response (1, 2). The TCR α and β chains are each composed of two immunoglobulin-like domains; most of the amino acid residues that are found to be highly conserved in the variable (V) region of immunoglobulins are also found in TCR V regions, suggesting that the tertiary structure of the TCR may resemble that of immunoglobulins (3, 4). However, TCR V regions have significantly more primary sequence variability, an increased apparent rate of divergence in phylogeny, and peaks of variability in addition to those noted in immunoglobulins (5, 6). In order to understand and control the molecular interactions underlying T-cell recognition of MHC/peptide complexes, complete structural knowledge of the TCR is required.

Several approaches have been employed to produce soluble, recombinant TCRs. In these recombinant TCR molecules, the transmembrane/cytoplasmic regions of α and β chains were replaced with sequences from lipid-linked proteins (7), the CD3 ζ chain (8), or immunoglobulins (9, 10). Soluble TCRs were either recovered as secreted proteins or obtained by enzymatic cleavage of the surface-expressed recombinant proteins. All of these approaches rely, however, on the assembly of the heterodimer, which is inefficient (11). In addition, high-level expression of the human TCR α chain in transfected eukaryotic cells is not stable. These problems can be avoided by the design of a single-chain (sc) recombinant protein in which the V regions of the heterodimer are joined by a short peptide linker. Such a design has been successfully applied to antibody molecules (12). Such recombinant molecules, scFv, have a specificity and affinity similar to that of native antibodies (12). Several reports have described the production of scTCRs in bacterial expression systems using the sc antibody (Fv) design (13–16), but none have presented functional data indicating that these scTCRs could recognize their MHC/peptide complexes or superantigens. Recently, however, the production in bacteria of a scTCR with the two-domain Fv design that could recognize its natural ligands has been reported, although the fraction correctly refolded was extremely low (17).

In the present report, different scTCR designs were evaluated in transfected eukaryotic cells with respect to surface expression of TCR molecules, proper folding, and recognition of the appropriate MHC/peptide ligand. A three-domain sc construct [α-chain V (V_α)–linker–β-chain V (V_β)–β-chain constant (C_β)] was stably expressed on the cell surface when linked to a glycosyl phosphatidylinositol (GPI) anchor and recognized by a conformation-dependent monoclonal antibody (mAb) specific for the $V_\beta17$ segment. The soluble form of this recombinant protein could be readily obtained by enzymatic cleavage with phosphatidylinositol-specific phospholipase C (PI-PLC). Replacement of the GPI domain with the cytoplasmic portion of the ζ chain resulted in a functional TCR molecule that transduced an intracellular signal following recognition of either the proper MHC/peptide or the MHC/staphylococcal enterotoxin B (SEB) superantigen complex. The production of a functional scTCR directly demonstrates the feasibility of employing sc design to produce soluble TCRs.

MATERIALS AND METHODS

Construction of Recombinant TCR Molecules. cDNAs of TCR α and β chains were prepared from mRNA of Hy.2H9 cells (18) with Superscript reverse transcriptase (BRL) and an oligo(dT) primer (Sigma) and were amplified by PCR using Vent DNA polymerase (New England Biolabs) and primers 5′-GCTCGAGGCGGCGATGGAAACTCTCCTGGGAGT-3′ (A5) and 5′-GGAATTCAGCTGGACCACAGCCGC-3′ for α-chain and 5′-GCTCGAGCTCTGCCATGGACTCCTG-GA-3′ and 5′-GGAATTCAGAAATCCTTTCTCTTGAC-3′ for β-chain. The cDNAs were cloned into the mammalian expression vector pBJ-neo (8). GPI-anchored TCR molecules (α-PI and β-PI) were constructed as follows. A *Ban* I site was

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TCR, T-cell receptor; MHC, major histocompatibility complex; sc, single chain; GPI, glycosyl phosphatidylinositol; PI-PLC, phosphatidylinositol-specific phospholipase C; SEB, staphylococcal enterotoxin B; MBP, myelin basic protein; V, variable; C, constant; IL-2, interleukin 2; mAb, monoclonal antibody.

‡Present address: Cystar, Inc., 30 Commerce Way, Woburn, MA 01801.

engineered after the fifth amino acid residue beyond the last cysteine by oligonucleotide-directed mutagenesis. The region 3′ of the *Ban* I site was then replaced with a *Ban* I–*Not* I fragment encoding the GPI signal domain from the human placental alkaline phosphatase. For the construction of various scTCRs, variable domains of the 2H9 TCR α and β chains were prepared by PCR using primers A5 and 5′-CAGAGCTCACGGATGAACAATAAGGCTGGT-3′ for the V_α domain in all the scTCR constructs, 5′-TCGGATATCGATGGTGGAATCACTCAGTCC-3′ (B5) and 5′-CAGAGATCAGCACGGTGAGCCGGTTCCC-3′ for the V_β domain in AB-PI-1, 5′-GTGGGAGATCTCTGCTTCTGATGGCTCAAAC and B5 for the V_β domain in AB-PI-2, 5′-CACGGATCCCCGTCTGCTCTACCCCAGGC and B5 for the V_β and C_β domains in ABC-PI, and 5′-CACGGATCCCCGTCTGCTCTACCCCAGGC-3′ and B5 for the V_β and C_β domains in ABC-ζ. The cDNA encoding the transmembrane and cytoplasmic domains of murine CD3 ζ chain (8) was a gift of R. D. Klausner (National Institutes of Health). Convenient restriction sites were engineered at the end of each fragment to aid in the assembly of the construct. The linker was a 15-amino acid motif of GGGGS repeated three times (12) with *Sac* I at the 5′ end and *Eco*RV at the 3′ end. Except for α-PI, all the constructs were cloned into pBJ-neo, which carries the G418-resistance gene. α-PI was cloned into pCEP-4 (Invitrogen), which bears the hygromycin-resistance gene. All constructs were verified by multiple restriction digests and by sequencing with the Sequenase kit (United States Biochemical).

Affinity Purification and Characterization of a Soluble Three-Domain scTCR. After transfection and G418 selection (8), cells expressing a high level of GPI-linked three-domain scTCR (ABC-PI) were isolated by three rounds of sorting. The resulting cells were grown in spinner culture to a density of 10^6 per ml and harvested by centrifugation. The pellet was washed twice with phosphate-buffered saline (PBS) and resuspended in PBS containing 2 mM Pefabloc (Centerchem, Stamford, CT) to a density of 5×10^7 per ml with PI-PLC (Sigma) added at 1 unit/ml. Cells were incubated at 37°C for 1 hr with constant rocking. The supernatant was collected by centrifugation and by passage through a 0.45-μm filter and applied to a column of Acti-gel (Sterogen, Arcadia, CA) with immobilized βF1. The column was washed with 10 volumes of PBS and the soluble TCR was eluted with 0.15 M glycine (pH 2.8). Fractions were immediately neutralized with 0.1 volume of saturated Tris. The soluble TCR was then dialyzed against >100 volumes of PBS at 4°C with at least four changes and concentrated to 0.5 mg/ml by vacuum dialysis against PBS. Five micrograms of purified soluble three-domain scTCR was analyzed by SDS/PAGE under reducing conditions.

Stimulation of Transfectants with Antibodies, SEB, and Peptide/MHC Ligands. ABC-ζ-transfected BW5147β $\alpha^-\beta^-$ (19) cells (5×10^4 per well) were cultured in a 96-well round-bottom plate to which various antibodies had been immobilized (1 μg per well). The supernatants were collected after 24 hr and interleukin 2 (IL-2) production was assessed in a bioassay using an IL-2-dependent cell line (CTLL) and the CellTiter-96 nonradioactive proliferation assay (Promega). In the case of ABC-ζ-transfected RBL-2H3 (8) cells, the cells were incubated with [^{3}H]serotonin (NEN) at 0.5 μCi (18.5 kBq) for 24 hr before they were added to the antibody plate. After incubation at 37°C for 2 hr, radioactivity released into the supernatant was measured in a liquid scintillation counter. The specific serotonin release was calculated as described (8). For SEB stimulation, 5×10^4 transfected cells per well were cultured with various concentrations of SEB (Toxin Technology, Sarasota, FL) in the presence or absence of 2×10^5 B cells. For antigen presentation, 5×10^4 transfected cells per well were cocultured with 2×10^5 B cells which were incubated with or without the myelin basic protein (MBP)-(85–99) peptide for 3 hr before the experiment. The assays were conducted as described above.

RESULTS

mAb C1 Recognizes a Conformational Epitope of TCR. Recombinant TCR molecules were generated by using the TCR α- and β-chain sequences of the human MBP-specific T-cell clone Hy.2H9 (18). This clone TCR is composed of the V_α3.1 and V_β17.1 segments and is specific for the immunodominant MBP peptide MBP-(85–99) in the context of HLA-DR2 (DRA, DRB1*1602) (18). Usage of the V_β17.1 segment allowed the proper folding of recombinant TCRs to be probed with the superantigen SEB (20) and the mAb C1 (21). To confirm the V_β17 specificity of C1, the extracellular domains of TCR α and β chains of Hy.2H9 cells were fused to the C-terminal sequence from human placental alkaline phosphatase for GPI anchorage (Fig. 1) and the DNAs encoding the GPI-anchored β and α chains (β-PI and α-PI) were sequentially transfected by electroporation (8) into a TCR α- and β-chain-deficient murine lymphoma cell line, BW5147 $\alpha^-\beta^-$ (BW$^-$) (19). The surface expression of the GPI-anchored TCR chains was monitored by staining with mAbs αF1 (22), βF1 (23), and C1. αF1 and βF1 recognize nonconformational epitopes located in the C region of the TCR α and β chains, respectively. Surface expression of the GPI-anchored TCR β chain is independent of heterodimer formation and assembly of the CD3 complex (7). In the β-PI-transfected cells (Fig. 2, open curves), high-level expression of β-PI was confirmed by staining with βF1. Interestingly, there was little C1 staining of these transfectants. However, when a GPI-anchored 2H9 α chain was supertransfected into the β-PI transfectant (Fig. 2, shaded curves), C1 reactivity was greatly increased while the level of βF1 staining remained constant. Thus, the V_β17-specific C1 epitope is conformational and dependent on the proper pairing of TCR α and β chains and can therefore be used to assess the proper folding of recombinant TCRs bearing a V_β17 sequence.

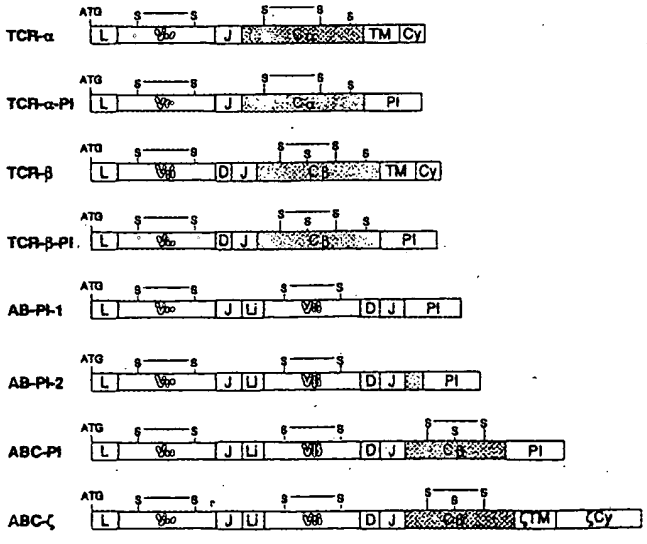


FIG. 1. Schematic representation of TCR α- and β-chain genes and various chimeric constructs. S—S, disulfide bond; L, leader; V, variable segment; J, joining segment; C, constant region; TM, transmembrane region; Cy, cytoplasmic region; ATG, start codon; Li, 15-residue peptide linker containing three repeats of GGGGS; PI, GPI domain of human placental alkaline phosphatase with the sequence LAPPAGTTDAAHPGRSVVPALLPLLAGTLLLL (7). The ζ region contains transmembrane and cytoplasmic domains of the murine CD3 ζ chain starting at position 31 (8).

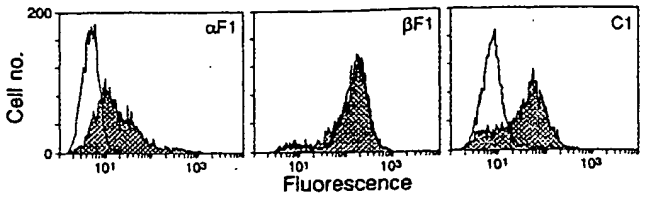


FIG. 2. Coexpression of α-PI and β-PI are required for C1 reactivity. Flow cytomeric analysis of β-PI-transfected BW$^-$ cells stained with mAb αF1, βF1, or C1 before (open curves) and after (shaded curves) the supertransfection of the α-PI construct.

High-level expression of the TCR α chain (α-PI) was, however, not stable either alone or in the presence of β-PI. Attempts were made on several cell lines, including COS-7, CHO-K1, and a TCR-deficient variant of Jurkat cells, JK-β^- (J.RT3-T3.5, American Type Culture Collection). The expression level of α-PI was comparable to that of β-PI after the initial drug selection, but continued culture for less than a month yielded a population of cells with little surface expression of α-PI, whereas β-PI expression was stable (data not shown). The inability to obtain cell lines with stable high-level expression of the PI-anchored human TCR α-chain has been reported by other laboratories as well (24).

Expression and Purification of a Three-Domain scTCR. To overcome the limitations set by the unstable expression of the human TCR α chain, various sc designs were examined. Initially, a design similar to that of sc antibodies (Fv) was chosen (12). A 15-residue flexible linker was used to link the C terminus of the V_α–J_α domain to the N terminus of the β chain. The GPI domain was then ligated to the C terminus of the V_β–J_β domain. The construct (AB-PI-1, Fig. 1) was transfected into several cell lines, including JK-β^-, COS-7, CHO-K1, and BW$^-$. Although the expression of the gene was confirmed by the detection of the correct RNA transcripts (Fig. 3*B*), no surface expression was detected, as evidenced by negative C1 antibody staining (Fig. 3*A*). Immunoprecipitation after metabolic labeling failed to recover any C1-reactive sc molecules from these transfectants (data not shown). The inability to identify any C1-reactive protein could have been due to the design of this molecule, such as insufficient linker length between the extracellular domain and the GPI domain. To improve the accessibility of the sc construct, another two-domain scTCR was designed in which an extra 30-amino acid portion of the N terminus of the C_β domain was added as a hinge region. The transfectants of this construct (AB-PI-2, Fig. 1) were still not reactive with the C1 antibody (data not shown). Finally, the entire C_β domain was added to the sc construct. A complete C_β domain should provide enough distance for the V_α–V_β domains to be expressed on the cell surface and, more importantly, should allow surface expression to be monitored with another antibody, βF1 (23). This three-domain scTCR was constructed by extending the TCR β-chain sequences to the residue right before the last cysteine (the sixth cysteine), which was then fused to the GPI domain. The last cysteine was deleted to prevent dimerization between C_β domains. This construct (ABC-PI, Fig. 1) was transfected into BW$^-$ cells and surface expression was confirmed by staining with both βF1 and C1 (Fig. 3*C*, shaded curves). Both antibodies stained the cells with comparable efficiency, suggesting that most of the molecules were expressed in the correct conformation. Moreover, the molecule could be efficiently cleaved from the cell surface with PI-PLC (Fig. 3*C*, open curves).

Soluble three-domain scTCR was purified from transfectants after PI-PLC cleavage followed by affinity chromatography using the βF1 antibody. The purified three-domain scTCR appeared as multiple bands at 50–70 kDa after SDS/PAGE (Fig. 3*D*). The heterogeneity of scTCR is probably the

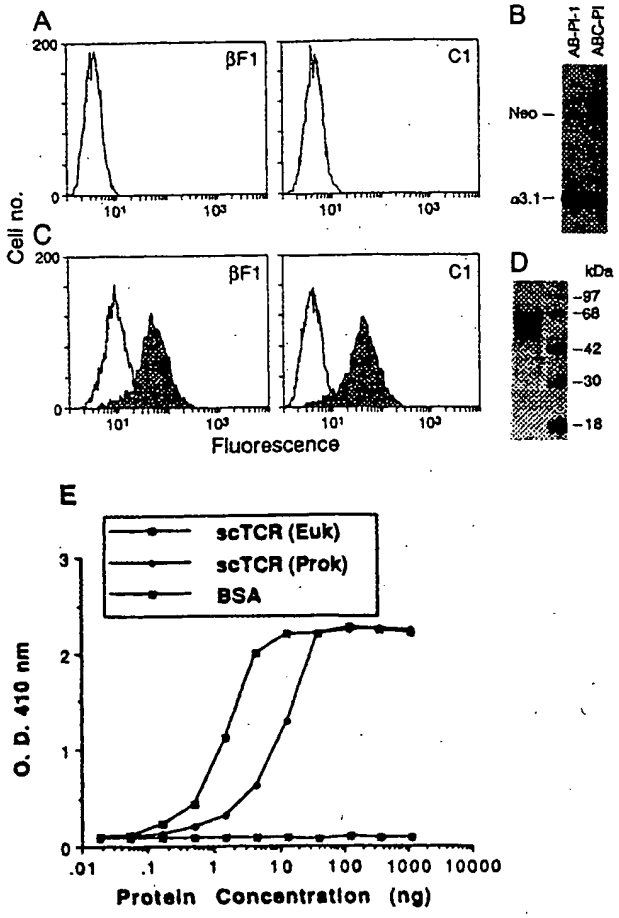


FIG. 3. Flow cytomeric analysis of surface expression of scTCR constructs. (*A*) Lack of C1 reactivity in BW$^-$ cells transfected with two-domain scTCR (AB-PI-1). (*B*) RNA analysis of poly(A)-enriched total cell RNA isolated from BW$^-$ cells transfected with two-domain and three-domain scTCR constructs AB-PI-1 and ABC-PI, respectively. Samples were analyzed in a nuclease S1 protection assay (25) using probes specific for the 5′ end of the transcripts from the TCR constructs. The coexpressed *neo* (G418-resistance gene) transcripts from the vector were analyzed with a probe at the same time as a control. (*C*) Flow cytometric analysis of ABC-PI-transfected BW$^-$ cells with both βF1 and C1 antibodies before (shaded curves) and after (open curves) PI-PLC treatment. (*D*) SDS/PAGE of affinity-purified three-domain scTCR. (*E*) Comparison of C1 reactivity of three-domain scTCRs produced from eukaryotic (Euk) and prokaryotic (Prok) expression systems in a two-antibody ELISA. A plateau is reached because the amount of βF1 attached to the plate became limiting.

result of variable glycosylation; its polypeptide size calculated from amino acid composition is 40 kDa. The structural integrity of the three-domain scTCR was verified by a two-antibody ELISA (Fig. 3*E*). The molecules were first captured by the βF1 antibody immobilized to the plate and then assessed for reactivity with the C1 antibody. When compared with the three-domain scTCR produced in a bacterial expression system (unpublished work), the scTCR from the eukaryotic system gave 10–20 times higher C1 reactivity. The purified three-domain scTCR was stable and could be stored in PBS at 4°C for months without significant loss of C1 reactivity.

Functional Characterization of a Chimeric Three-Domain scTCR. To directly assess the functional integrity of the three-domain scTCR, a self-signaling scTCR was produced by replacing the GPI domain with the transmembrane and cytoplasmic domains of the CD3 ζ chain. These regions have been shown to be sufficient for signal transduction when its

extracellular fusion partner is crosslinked by an antibody or by the proper ligand (8, 26, 27). To enable the recovery of three-domain scTCR as a soluble form, a linker containing a thrombin cleavage site was inserted into the junction of three-domain scTCR and the ζ domain. The construct (ABC-ζ) was transfected into BW$^-$ cells (28) and the rat basophilic leukemia cell line RBL-2H3 (RBL) (8), and the populations displaying high-level expression of three-domain scTCR were isolated by three rounds of cytofluorometric sorting using the antibody βF1. The ABC-ζ-transfected cells were first stimulated with various antibodies to confirm the self-signaling nature of this recombinant molecule. The signal transduced upon the activation of the three-domain scTCR was measured as IL-2 production in BW$^-$ transfectants, whereas serotonin release was measured in RBL transfectants. Both transfectants showed a strong response following βF1 and C1 stimulation but not to purified mouse immunoglobulin or anti-CD8 antibody used as controls (Fig. 4 *A* and *B*). The structural integrity of the scTCR was further examined with the superantigen SEB, which binds to both V$_\beta$17 and MHC class II molecules, resulting in TCR crosslinking and T-cell activation regardless of the peptide bound to the MHC molecule (20, 29). ABC-ζ transfectants displayed a concentration-dependent response toward SEB (Fig. 4*C*) when the superantigen was presented by transformed B-cell lines with high-level expression of DR1 (DRA, DRB1*0101; cell line LG2) or DR2 (DRA, DRB1*1602; cell line 9016). Thus, the lateral face of the TCR V$_\beta$ region to which SEB is thought to bind (30) is structurally intact.

To prove that the three-domain scTCR did indeed recognize the MHC/peptide ligand, antigen presentation experiments using the natural ligand for the Hy.2H9 clone, 9016 cells bearing the DRB1*1602 allele of DR2, and MBP-(85–99) peptide were performed. To ensure detection of subtle abnormalities in the structure of the three-domain scTCR, 9009 cells (DRA, DRB1*1601), which also bind MBP-(85–99), were used as a control. DRB1*1601 and DRB1*1602 differ only at position 67 in the DR$_\beta$1 domain; this TCR contact-residue substitution does, however, abolish recognition of the peptide by the parent T-cell clone (ref. 18; K.W.W.,

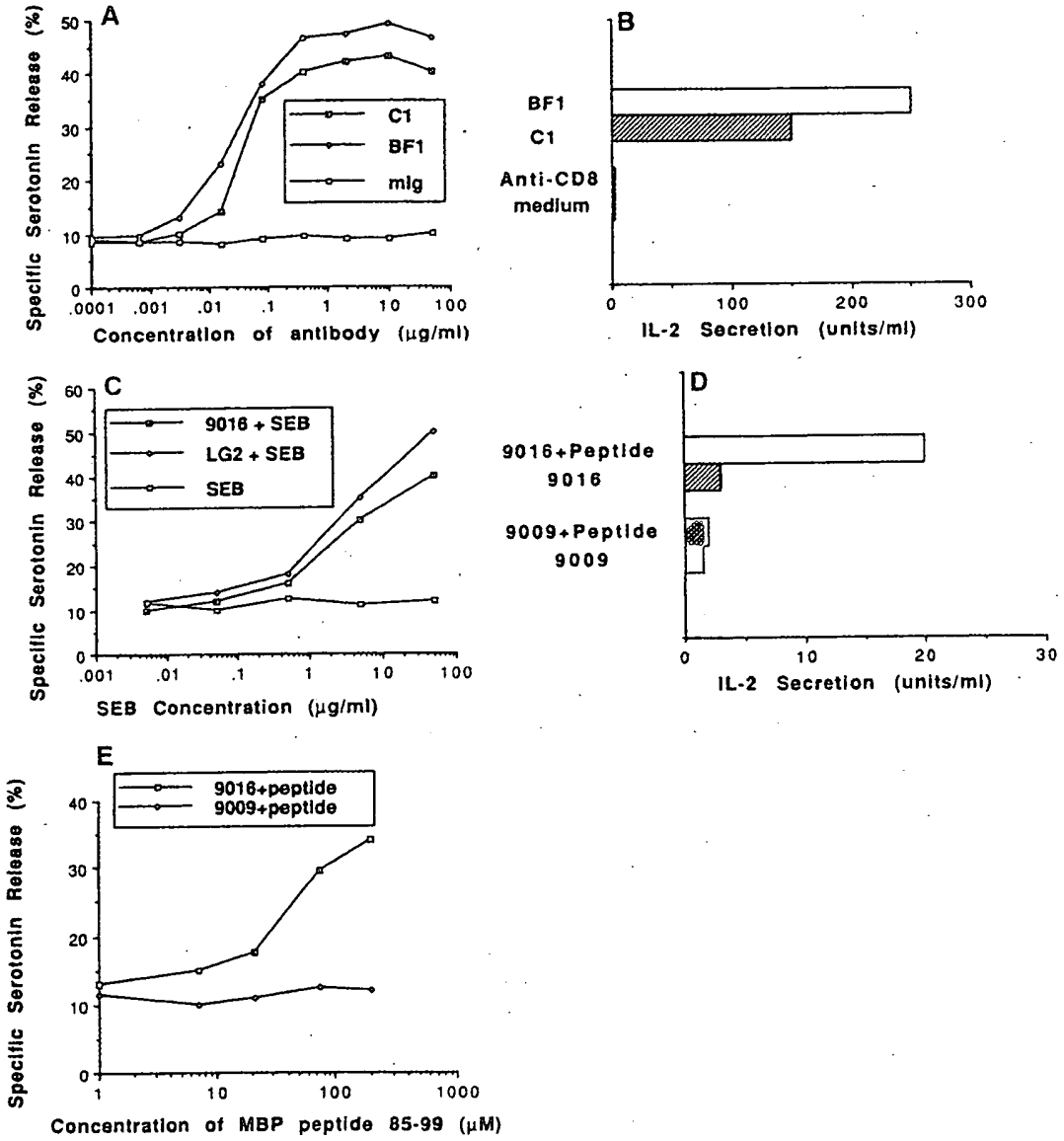


FIG. 4. (*A*) [^{3}H]Serotonin release following TCR crosslinking of three-domain scTCR–ζ chimeric construct (ABC-ζ)-transfected RBL cells. C1 is specific to V$_\beta$17 and βF1 is specific to the C region of TCR β chain. mIg, mouse immunoglobulin. (*B*) IL-2 release following TCR crosslinking of ABC-ζ-transfected BW$^-$ cells. (*C*) [^{3}H]Serotonin release following SEB stimulation of ABC-ζ-transfected RBL cells. (*D*) MHC-restricted antigen-specific response from the ABC-ζ-transfected BW$^-$ cells. (*E*) Dose–response curve of MHC/peptide recognition of ABC-ζ-transfected RBL cells.

unpublished work). The ABC-ζ-transfected BW$^-$ cells secreted IL-2 in response to peptide-pulsed 9016 cells, but not to peptide-pulsed 9009 cells (Fig. 4*D*). Similar results were obtained with RBL transfectants (Fig. 4*E*), as serotonin release was dependent on the concentration of the MBP peptide used to pulse 9016 cells. The signal appeared to be weak when compared with antibody stimulation. This is not surprising, however, since saturating amounts of antibodies are expected to crosslink the majority of TCR molecules on the target cell, whereas a much smaller fraction of TCR molecules is probably engaged when T cells are cocultured with peptide-pulsed antigen-presenting cells, in which a maximum of 5–15% of the DR molecules bind the peptide. The requirement for a high concentration of peptide or SEB is not due to the sc design, since high concentrations of peptide are also needed to stimulate $\alpha\beta$ heterodimers of the TCR-ζ constructs (8). It is likely that the decrease in sensitivity results from the lack of CD3, CD4, and/or other adhesion/signaling molecules. Nonetheless, these results demonstrate that the three-domain scTCR was correctly folded and functionally competent. In addition, a soluble form of three-domain scTCR could be obtained from the ABC-ζ transfectants by thrombin cleavage and affinity purification (data not shown).

DISCUSSION

A scTCR molecule was designed which contains the V domains of both α and β chains and the C domain of the β chain. This scTCR molecule could be stably expressed at a high level in eukaryotic cells and could be isolated in a soluble form by enzymatic cleavage and affinity chromatography. The V_α and V_β domains appeared to be properly paired, since the scTCR bound to a conformation-dependent mAb, the superantigen SEB, and the proper MHC/peptide ligand. This design of scTCR offers an alternative to the two-chain design of soluble TCRs and has several advantages. (*i*) The sc design avoids the low-efficiency dimerization process which may be the limiting step in the assembly of TCR heterodimers from α and β subunits made in *Escherichia coli*. The sc design therefore allows efficient expression of the recombinant protein in quantities suitable for structural analysis and for some diagnostic or therapeutic applications. (*ii*) The design avoids the problems associated with the unstable expression of the human TCR α chain that have hindered efficient expression of human TCR molecules in eukaryotic cells. (*iii*) The sc design may allow the construction of TCR phage display libraries similar to those made for sc antibodies (28, 31). scTCR phage libraries may be powerful tools for the isolation of TCRs with defined specificities and/or high affinity for selective targeting of malignant and virally infected cells and for analyzing the interactions among TCR, MHC/peptide complexes, and superantigens.

Unlike antibodies, separately expressed V domains of TCR α and β chains have not been reconstituted to form heterodimers (ref. 32; K. L. Hilyard, personal communication). However, despite this success in producing the three-domain scTCR, a two-domain scTCR with detectable C1 reactivity could not be produced either in eukaryotic cells or in bacteria. The presence of the C_β domain (or part of the domain) may be required for the proper folding and/or stabilization of a scTCR molecule. In any event, the three-domain design provides a general means for the efficient production of functional scTCR molecules.

We thank Drs. James Huston and Herman Oppermann for helpful advice and Dr. Basya Rybalov for technical assistance. This work was supported by National Institutes of Health Grant CA47554 (J.L.S.), and National Multiple Sclerosis Society Postdoctoral Fellowship Award FG915-A-1 (S.C.). K.W.W. is a Harry Weaver Neuroscience Scholar of the National Multiple Sclerosis Society.

1. Ashwell, J. D. (1990) *Annu. Rev. Immunol.* **8**, 139–167.
2. Jorgensen, J. L., Reay, P. A., Ehrich, E. W. & Davis, M. M. (1992) *Annu. Rev. Immunol.* **10**, 835–837.
3. Davis, M. M. & Bjorkman, P. J. (1988) *Nature (London)* **334**, 395–402.
4. Chothia, C., Boswell, D. R. & Lesk, A. M. (1988) *EMBO J.* **7**, 3745–3755.
5. Behlke, M. A., Spinella, D. G., Chou, H. S., Sha, W., Hartl, D. L. & Loh, D. Y. (1985) *Science* **229**, 566–570.
6. Becker, D. M., Patten, P., Chien, Y.-H., Yokota, T., Eshaar, Z., Giedlin, M., Gascoigne, N. R. J., Goodnow, C., Wolf, R., Arai, A. & Davis, M. M. (1985) *Nature (London)* **317**, 430–434.
7. Lin, A. Y., Devaux, B., Green, A., Sagerstrom, C., Elliott, J. F. & Davis, M. M. (1990) *Science* **249**, 677–679.
8. Engel, I., Ottenhoff, T. H. M. & Klausner, R. D. (1992) *Science* **256**, 1318–1321.
9. Weber, S., Traunecker, A., Oloveri, F., Gerhard, W. & Karjalainen, K. (1992) *Nature (London)* **356**, 793–796.
10. Gregoire, C., Rebai, N., Schweisguth, F., Necker, A., Mazza, G., Auphan, N., Millward, A., Schmitt-Verhulst, A.-M. & Malissen, B. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8077–8081.
11. Klausner, R. D., Lippincott-Schwartz, J. & Bonifacino, J. S. (1990) *Annu. Rev. Cell Biol.* **6**, 403–431.
12. Huston, J., McCartney, J., Tai, M.-S., Mottola-Hartshorn, C., Jin, D., Warren, F., Keck, P. & Oppermann, H. (1992) *Int. Rev. Immunol.* **10**, 195–217.
13. Novotny, J., Ganju, R. K., Smiley, S. T., Hussey, R. E., Luther, M. A., Recny, M. A., Siliciano, R. F. & Reinherz, E. L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8646–8650.
14. Ward, E. S. (1992) *J. Mol. Biol.* **224**, 885–890.
15. Soo Hoo, W. F., Lacy, M. J., Denzin, L. K., Voss, E. W., Jr., Hardman, K. D. & Kranz, D. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4759–4763.
16. Kurucz, I., Jost, C. R., George, A. J. T., Andrew, S. M. & Segal, D. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3830–3834.
17. Hilyard, K. L., Reyburn, H., Chung, S., Garboczi, D. N., Bell, J. I. & Strominger, J. L. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9057–9061.
18. Wucherpfennig, K. W., Zhang, J., Witek, C., Matsui, M., Modabber, Y., Ota, K. & Hafler, D. A. (1994) *J. Immunol.* **152**, 5581–5592.
19. White, J., Blackman, M., Bill, J., Kappler, J., Marrack, P., Gold, D. & Born, W. (1989) *J. Immunol.* **143**, 1822–1828.
20. Choi, Y., Kotzin, B., Herron, L., Callahan, J., Marrack, P. & Kappler, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8941–8945.
21. Friedman, S. M., Crow, M. K., Tumang, J. R., Tumang, M., Xu, Y., Hodtsev, A. S., Cole, B. C. & Posnett, D. N. (1991) *J. Exp. Med.* **174**, 891–900.
22. Henry, L., Tian, W. T., Rittershaus, C., Ko, J. L., Marsh, H. C., Jr., & Ip, S. H. (1990) *Hybridoma* **8**, 577–588.
23. Brenner, M. B., Mclean, J., Scheft, H., Warnke, R. A., Jones, N. & Strominger, J. L. (1987) *J. Immunol.* **138**, 1502–1509.
24. Devaux, B., Bjorkman, P. J., Stevenson, C., Greil, W., Elliott, J. F., Sagerstrom, C., Clayberger, C., Krensky, A. M. & Davis, M. M. (1991) *Eur. J. Immunol.* **21**, 2111–2119.
25. Chung, S. & Perry, R. P. (1993) *Nucleic Acids Res.* **21**, 3301–3308.
26. Irving, B. A. & Weiss, A. (1991) *Cell* **64**, 891–898.
27. Romeo, C. & Seed, B. (1991) *Cell* **64**, 1037–1045.
28. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991) *Science* **254**, 624–628.
29. Irwin, M. J., Hudson, K. P., Ames, K. T., Fraser, J. D. & Gascoigne, N. R. (1993) *Immunol. Rev.* **131**, 61–78.
30. Choi, Y., Herman, A., Digiusto, D., Wade, T., Marrack, P. & Kappler, J. (1990) *Nature (London)* **346**, 471–473.
31. McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990) *Nature (London)* **348**, 552–554.
32. Ward, E. S. (1991) *Scand. J. Immunol.* **34**, 215–220.

US005759817A

United States Patent [19]

Barbas

[11] Patent Number: 5,759,817

[45] Date of Patent: Jun. 2, 1998

[54] **HETERODIMERIC RECEPTOR LIBRARIES USING PHAGEMIDS**

[75] Inventor: **Carlos Barbas**, San Diego, Calif.

[73] Assignee: **The Scripps Research Institute**, La Jolla, Calif.

[21] Appl. No.: **322,730**

[22] Filed: **Oct. 12, 1994**

Related U.S. Application Data

[63] Continuation of Ser. No. 826,623, Jan. 27, 1992, abandoned, which is a continuation-in-part of Ser. No. 683,602, Apr. 10, 1991, abandoned.

[51] **Int. Cl.**[6] **G01N 33/53**; C12N 15/00; C12P 21/00; C07K 16/00

[52] **U.S. Cl.** **435/69.7**; 435/7.1; 435/172.3; 435/320.1; 530/387.1; 530/387.3

[58] **Field of Search** ..... 69/1 T; 435/172.2, 435/5, 7.1, 172.1, 172.3, 320.1, 69.7; 536/24; 935/8; 530/387.1, 387.3, 412

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,223,409	6/1993	Ladner	435/69.7
5,403,484	4/1995	Ladner	435/235.1
5,427,908	6/1995	Dower et al.	435/5

FOREIGN PATENT DOCUMENTS

0 368 684	5/1990	European Pat. Off. .
WO91/17271	11/1991	WIPO .
WO91/18980	12/1991	WIPO .
WO91/19818	12/1991	WIPO .
WO92/01047	1/1992	WIPO .

OTHER PUBLICATIONS

Gussow, et al., *Cold Spring Harbor Symposia on Quantitative Biology*, vol. LIV, pp. 265–272 (1989).

Bass et al., *Proteins: Structure, Function and Genetics*, 8:309–314 (1990).

Bird et al., *Science*, 242:423–426 (1988).

Boss et al., *Nucl. Acids Res.*, 12:3791–3806 (1984).

Cabilly et al., *Proc. Natl. Acad. Sci.*, USA, 81:3273–3277 (1984).

Caton et al., *Proc. Natl. Acad. Sci., USA*, 87:6450–6454 (1990).

Cwirla et al., *Proc. Natl. Acad. Sci., USA*, 87:6378–6382 (1990).

Il'ichev et al., *Molekulyarnaya Biologiya*, 24:530–535 (1990).

Larrick et al., *Bio/Technology*, 7:934–938 (1989).

Model et al., "The Bacteriophages", ed: R. Calender, Plenum Publishing Corp., 2:375–456 (1988).

Mullinax et al., *Proc. Natl. Acad. Sci., USA*, 87:8095–8099 (1990).

Orlandi et al., Proc. Natl. Acad. Sci., USA, 86:3833–3837 (1989).

Parmely et al., Gene, 73:305–318 (1988).

Pluckthun, *Nature*, 347:497–498 (1990).

Rasched et al., *Microbiol. Rev.*, 50:401–427 (1986).

Sastry et al., *Proc. Natl. Acad. Sci.*, USA, 86:5728–5732 (1989).

Scott et al., *Science*, 249:386–390 (1990).

Short et al., Nuc. Acids Res., 16:7583–7600 (1988).

Skerra et al., *Science*, 240:1038–1041 (1988).

Skerra et al., *Bio/Technology*, 9:273–278 (1991).

Smith, G., *Science*, 228:1315–1317 (1985).

Takeda et al., *Nature*, 314:452–454 (1985).

Ward et al., Nature, 341:544–546 (1989).

Barbas et al., *Methods: A Companion to Methods in Enzymology*, 2:119–124 (1991).

Huse et al., *Science*, 246:1275–1281 (1989).

Kang et al., Proc. Natl. Acad. Sci., USA, 88:4363–4366 (1991).

McCafferty et al., *Nature*, 348:552–554 (1990).

Persson et al., *Proc. Natl. Acad. Sci., USA*, 88:2432–2436 (1991).

Barbas, C.F. and Lerner, R.A., "Combinational Immunoglobulin Libraries on the Surface of Phage (Phabs): Rapid Selection of Antigen-Specific Fabs" *Methods: A Companion to Methods in Enzymology*, 2:119–124 (1991).

Barbas, et al., "Assembly of Combinational Antibody Libraries on Phage Surfaces: The Gene III Site", *Proc. Natl. Acad. Sci. USA*, 88:7978–7982 (1991).

Kang, et al., "Linkage of Recognition and Replication Functions by Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces", *Proc. Natl. Acad. Sci. USA*, 88:4363–4366 (1991).

Higuchi 1989, in PCR Technology, Henry A. Erlich ed "M" Stockton Press.

Sambrook et. al. 1989 Molecular Cloning, pp. 15.51–15.84, Cold Springs Harbor Press, Cold Spring Harbor NY.

Primary Examiner—Nancy Degen
Assistant Examiner—Sean M. Garry
Attorney, Agent, or Firm—Thomas Fitting; Emily Holmes

[57] **ABSTRACT**

Filamentous phage comprising a matrix of cpVIII proteins encapsulating a genome encoding first and second polypeptides of an antogenously assembling receptor, such as an antibody, and a receptor comprised of the first and second polypeptides surface-integrated into the matrix via a cpVIII membrane anchor domain fused to at least one of the polypeptides with a mutagenized CDR3 region.

26 Claims, 12 Drawing Sheets

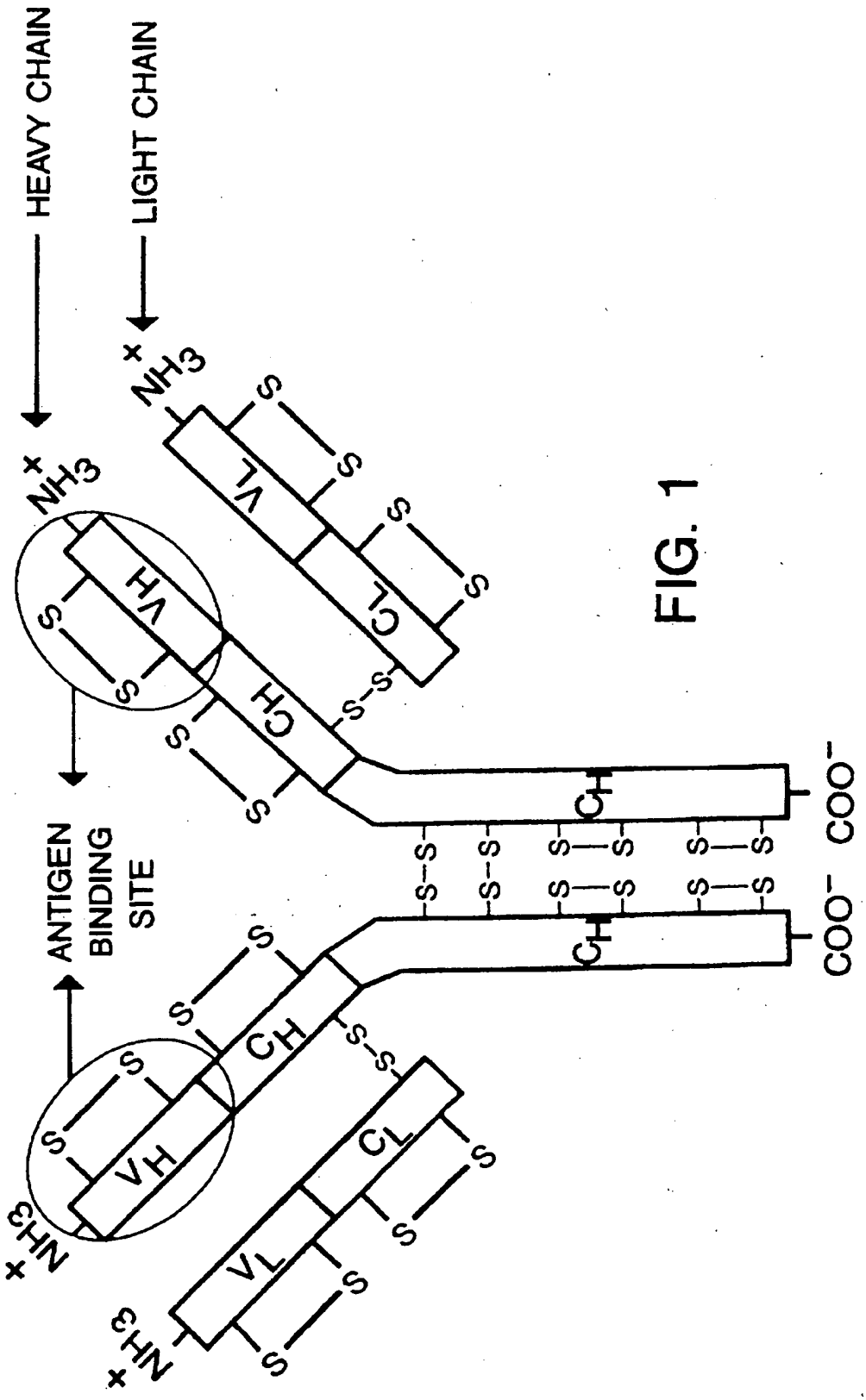

HEAVY CHAIN
LIGHT CHAIN
NH3
VL
CL
VH
CH
ANTIGEN BINDING SITE
COO-
S-S

FIG. 1

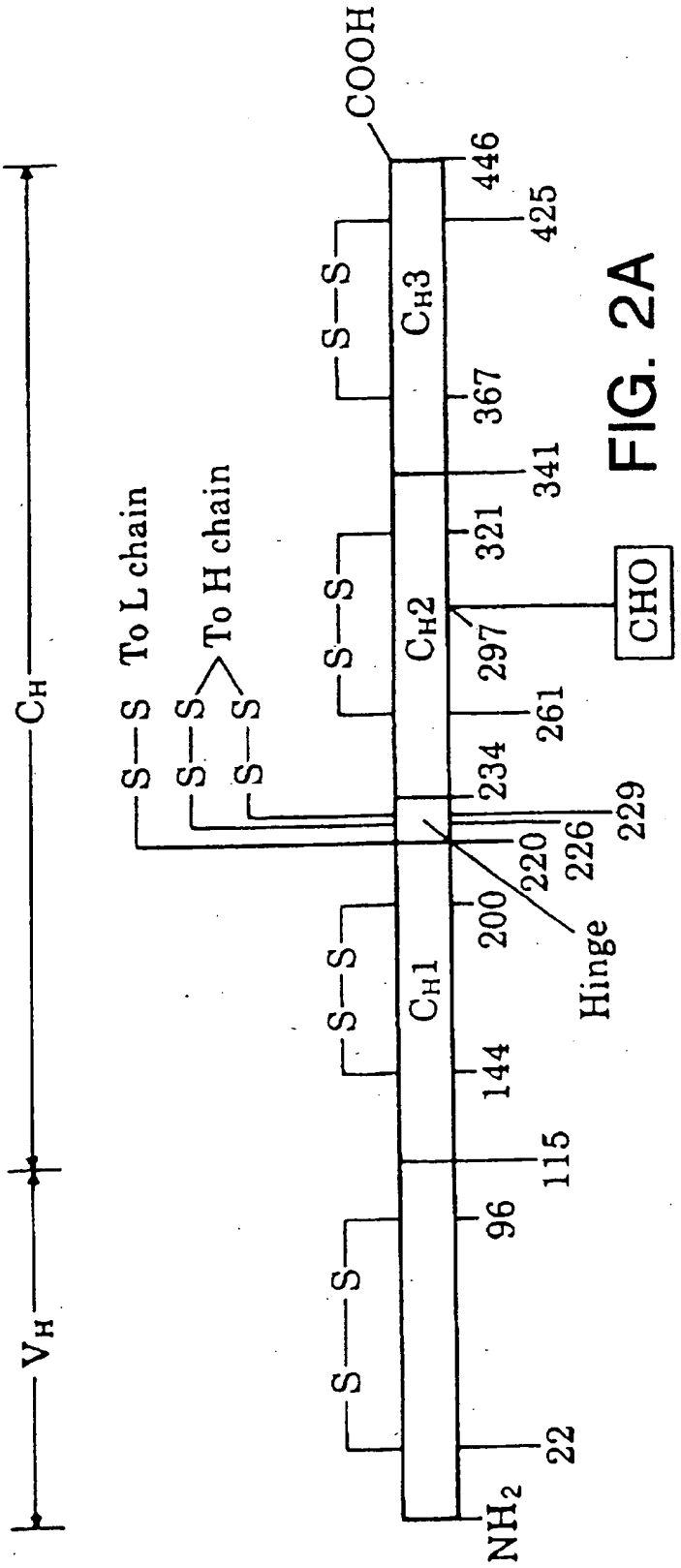

COOH
446
425
C_H3
367
341
321
CHO
C_H2
297
261
234
229
226
220
200
Hinge
C_H1
144
115
96
22
NH_2
S—S
To L chain
To H chain
C_H
V_H

FIG. 2A

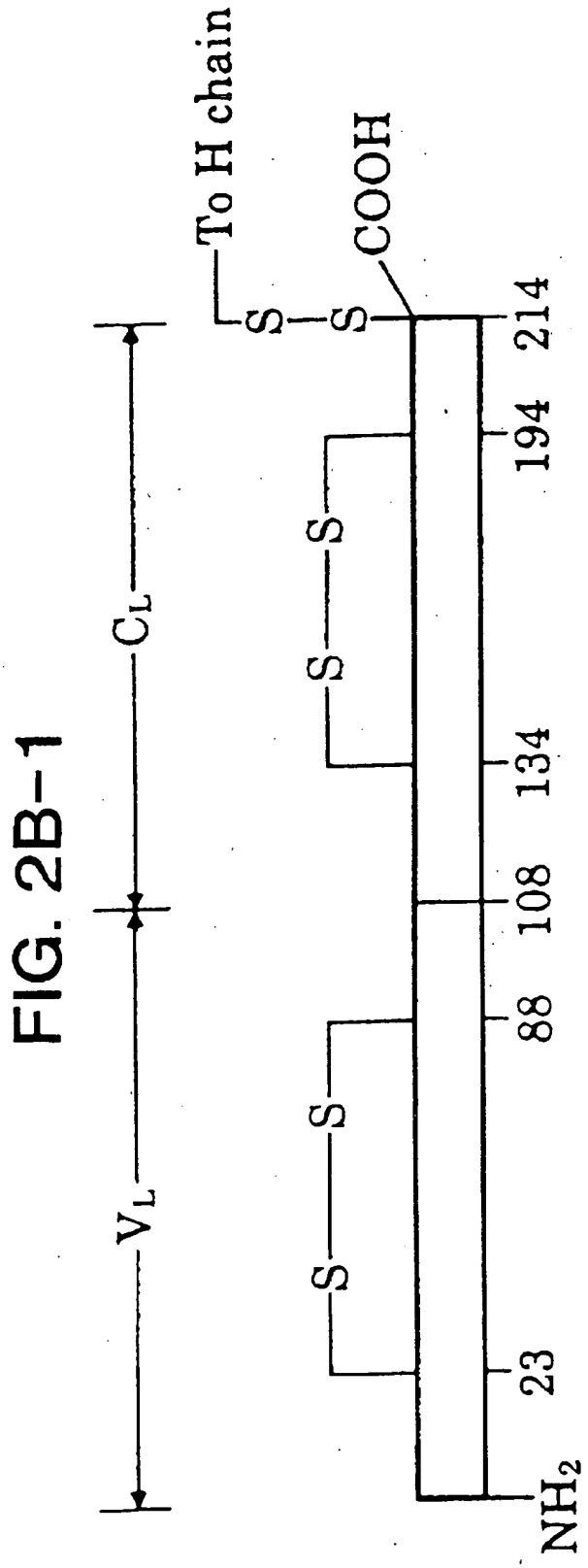


FIG. 2B-1

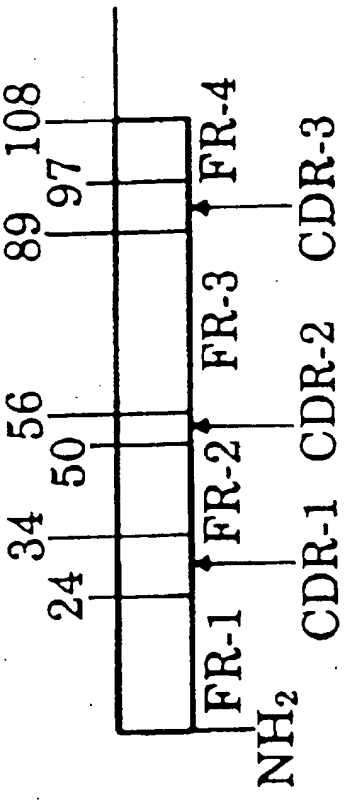


FIG. 2B-2

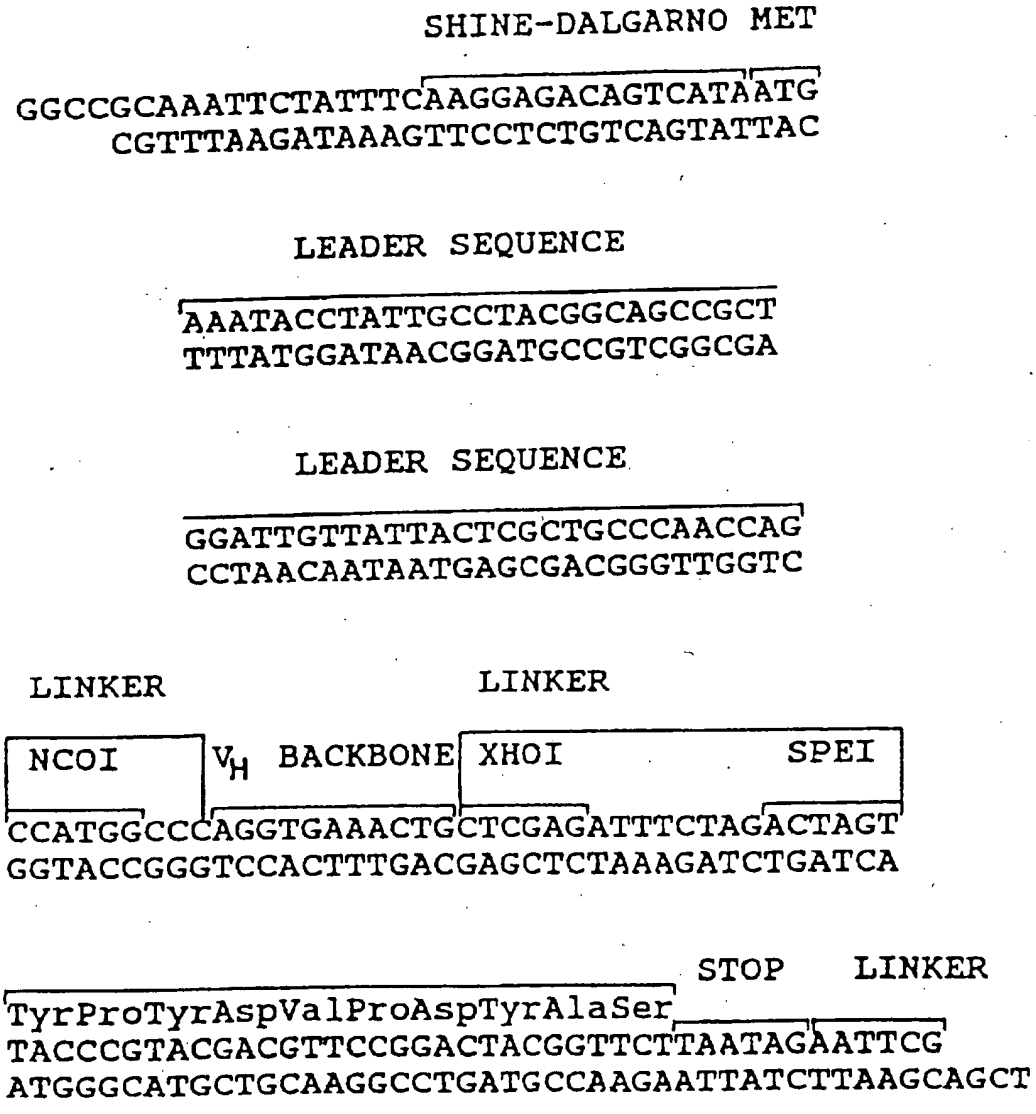

SHINE-DALGARNO MET
GGCCGCAAATTCTATTTCAAGGAGACAGTCATAATG
CGTTTAAGATAAAGTTCCTCTGTCAGTATTAC
LEADER SEQUENCE
AAATACCTATTGCCTACGGCAGCCGCT
TTTATGGATAACGGATGCCGTCGGCGA
LEADER SEQUENCE
GGATTGTTATTACTCGCTGCCCAACCAG
CCTAACAATAATGAGCGACGGGTTGGTC
LINKER
LINKER
NCOI
V_H BACKBONE
XHOI
SPEI
CCATGGCCCAGGTGAAACTGCTCGAGATTTCTAGACTAGT
GGTACCGGGTCCACTTTGACGAGCTCTAAAGATCTGATCA
STOP
LINKER
TyrProTyrAspValProAspTyrAlaSer
TACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTCG
ATGGGCATGCTGCAAGGCCTGATGCCAAGAATTATCTTAAGCAGCT

FIG. 3

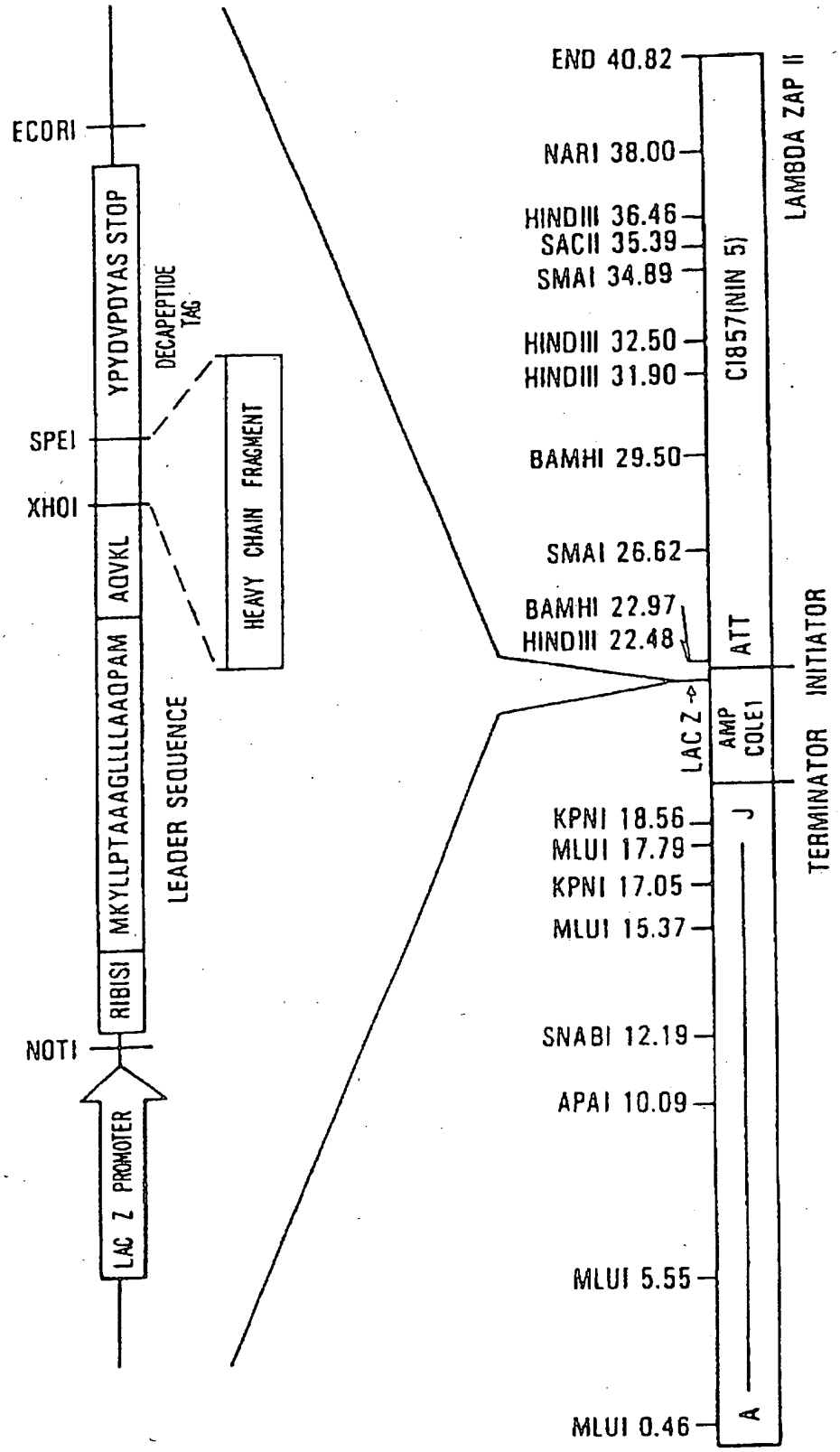

ECORI
YPYDVPDYAS STOP
DECAPEPTIDE TAG
SPEI
XHOI
AQVKL
HEAVY CHAIN FRAGMENT
MKYLLPTAAAGLLLAAQPAM
LEADER SEQUENCE
RIBISI
NOTI
LAC Z PROMOTER
END 40.82
NARI 38.00
HINDIII 36.46
SACII 35.39
SMAI 34.89
HINDIII 32.50
HINDIII 31.90
BAMHI 29.50
SMAI 26.62
BAMHI 22.97
HINDIII 22.48
C1857(NIN 5)
LAMBDA ZAP II
ATT
INITIATOR
LAC Z
AMP
COLE1
TERMINATOR
KPNI 18.56
MLUI 17.79
KPNI 17.05
MLUI 15.37
J
SNABI 12.19
APAI 10.09
MLUI 5.55
MLUI 0.46
A

FIG. 4

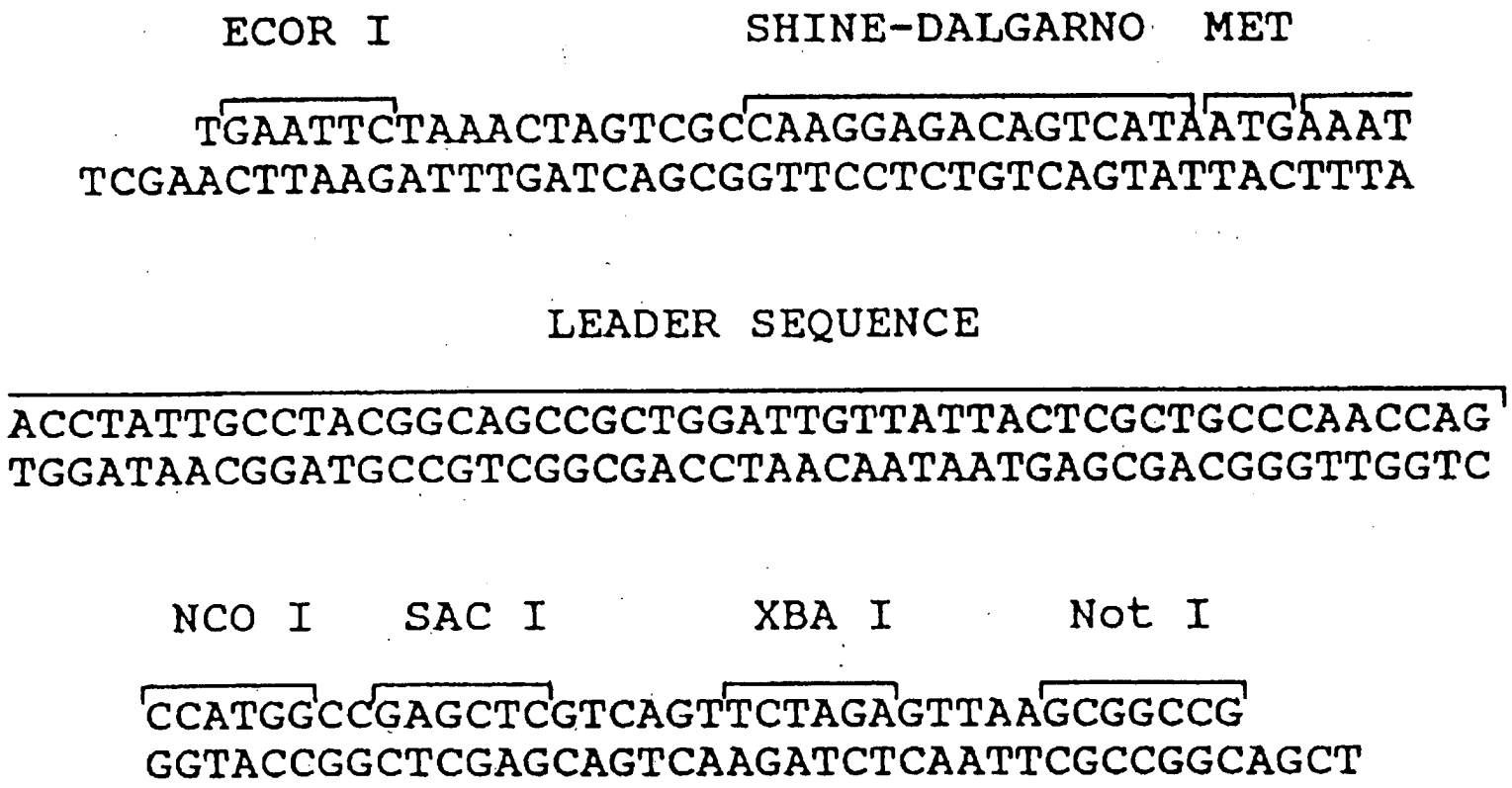

ECOR I
SHINE-DALGARNO
MET
TGAATTCTAAACTAGTCGCCAAGGAGACAGTCATAATGAAAT
TCGAACTTAAGATTTGATCAGCGGTTCCTCTGTCAGTATTACTTTA
LEADER SEQUENCE
ACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAG
TGGATAACGGATGCCGTCGGCGACCTAACAATAATGAGCGACGGGTTGGTC
NCO I
SAC I
XBA I
Not I
CCATGGCCGAGCTCGTCAGTTCTAGAGTTAAGCGGCCG
GGTACCGGCTCGAGCAGTCAAGATCTCAATTCGCCGGCAGCT

FIG. 5

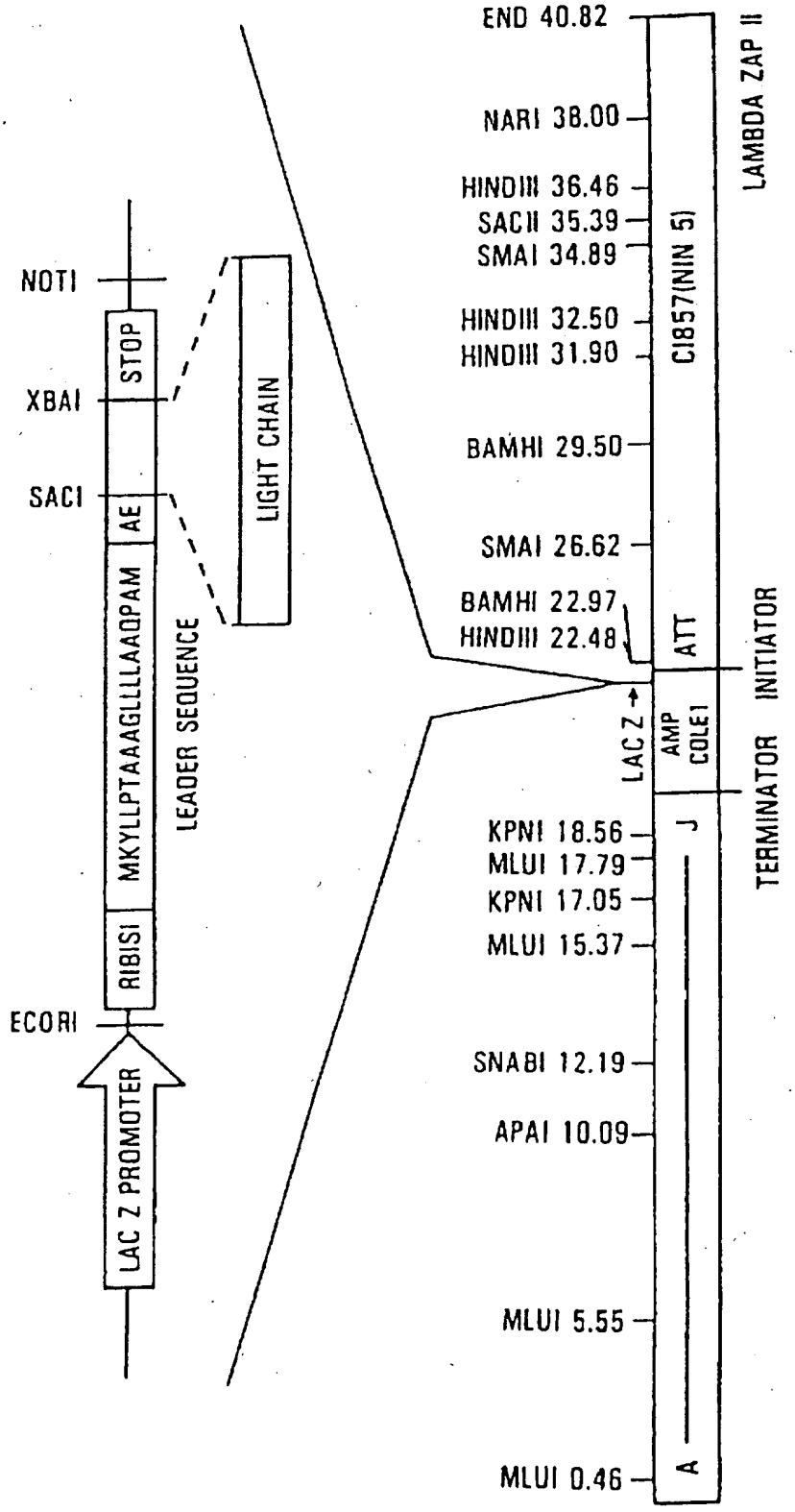

END 40.82
NARI 38.00
HINDIII 36.46
SACII 35.39
SMAI 34.89
HINDIII 32.50
HINDIII 31.90
BAMHI 29.50
SMAI 26.62
BAMHI 22.97
HINDIII 22.48
LAC Z
KPNI 18.56
MLUI 17.79
KPNI 17.05
MLUI 15.37
SNABI 12.19
APAI 10.09
MLUI 5.55
MLUI 0.46
LAMBDA ZAP II
CI857(NIN 5)
ATT
AMP
COLE1
J
A
TERMINATOR
INITIATOR
NOTI
XBAI
SACI
ECORI
STOP
AE
RIBISI
MKYLLPTAAAGLLLAAQPAM
LEADER SEQUENCE
LIGHT CHAIN
LAC Z PROMOTER

FIG. 6

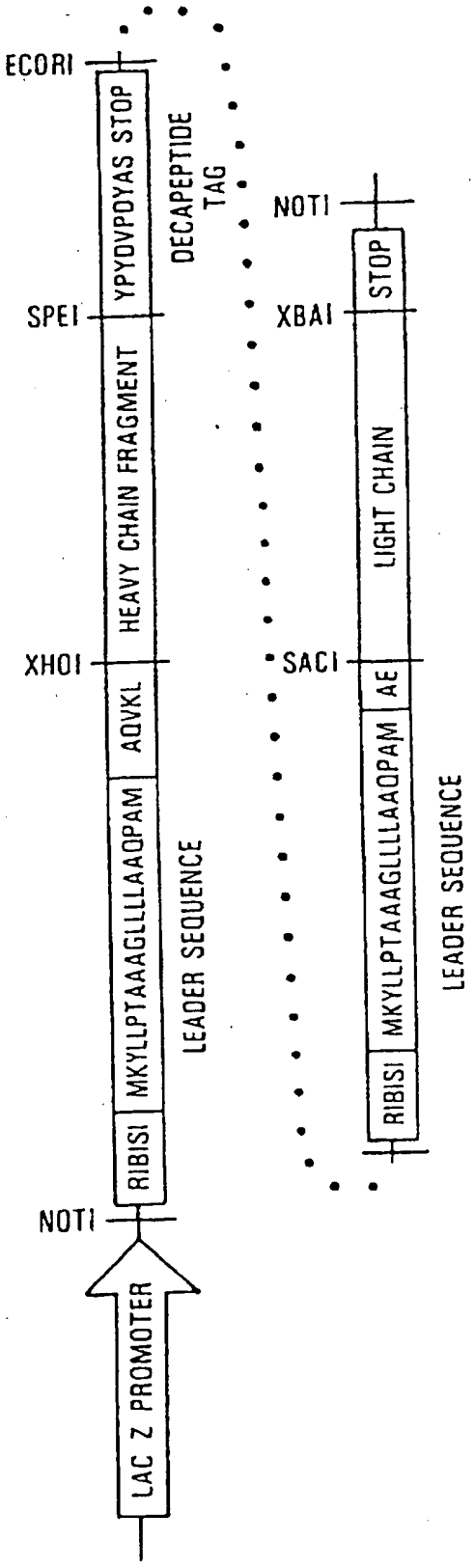

ECORI
SPEI
XHOI
NOTI
LAC Z PROMOTER
RIBISI
MKYLLPTAAAGLLLLAAQPAM
AQVKL
LEADER SEQUENCE
HEAVY CHAIN FRAGMENT
YPYDVPDYAS STOP
DECAPEPTIDE TAG
NOTI
XBAI
SACI
RIBISI
MKYLLPTAAAGLLLLAAQPAM
AE
LEADER SEQUENCE
LIGHT CHAIN
STOP

FIG. 7

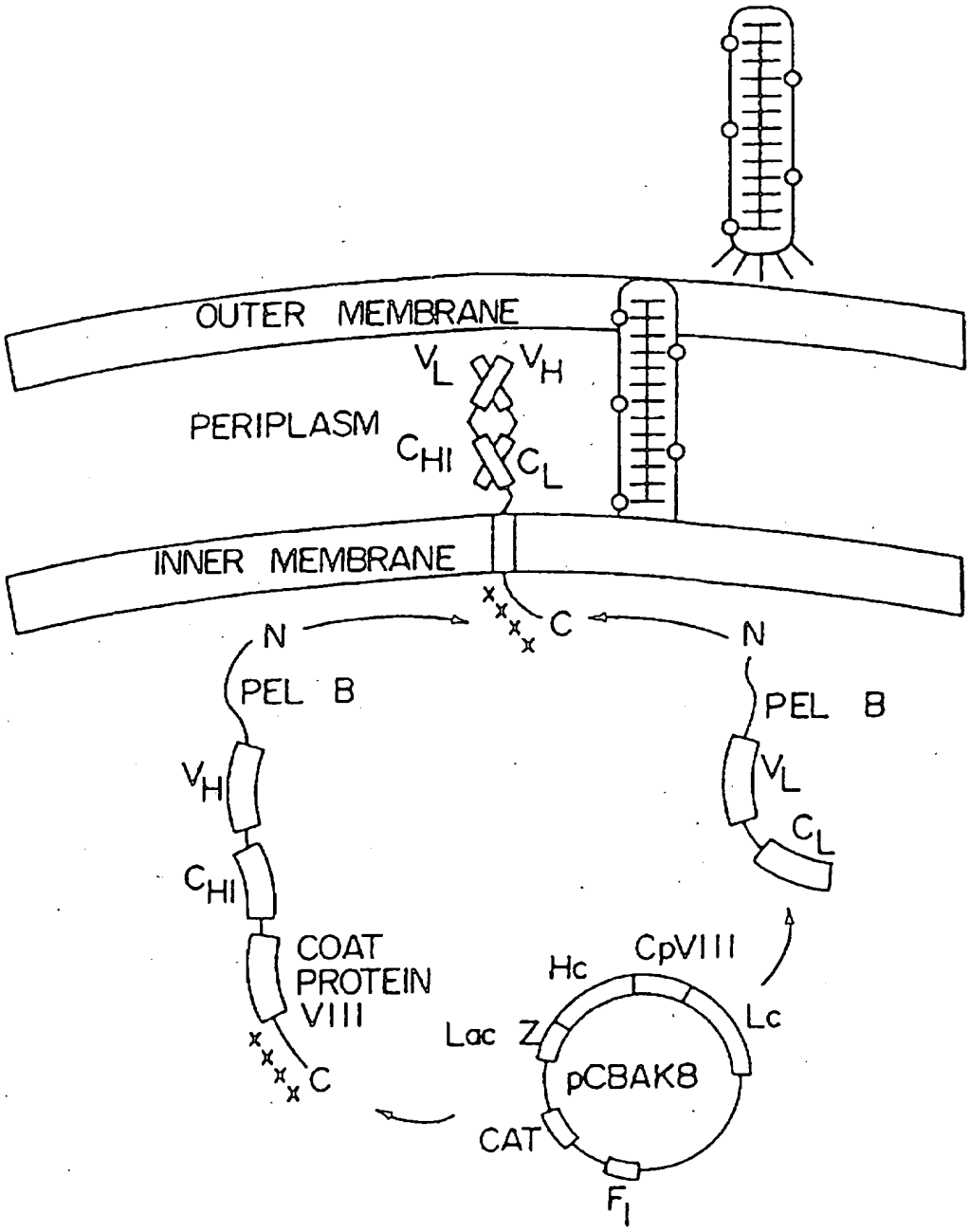

OUTER MEMBRANE
V_L
V_H
PERIPLASM
C_{HI}
C_L
INNER MEMBRANE
N
C
N
PEL B
PEL B
V_H
V_L
C_{HI}
C_L
COAT PROTEIN VIII
C
Hc
CpVIII
Lac Z
Lc
pCBAK8
CAT
F_1

FIG. 8

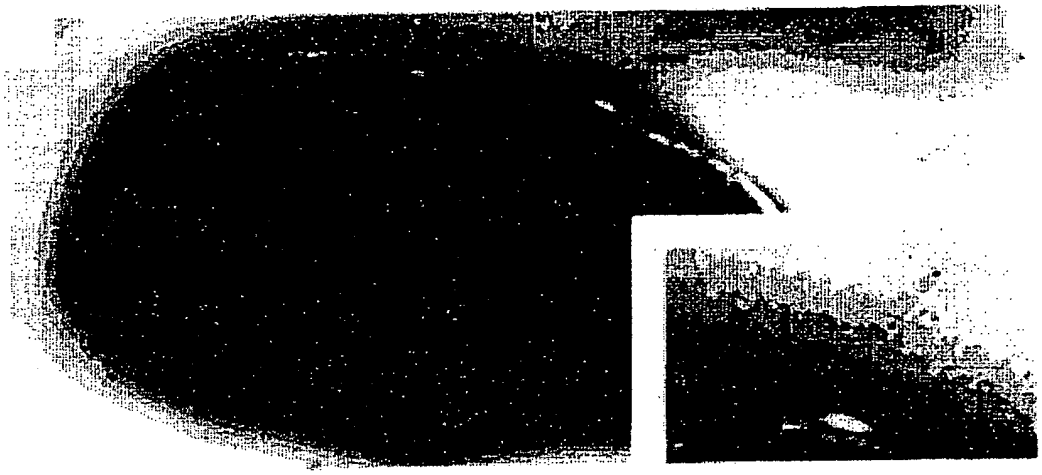


FIG. 9A-1

FIG. 9A-2

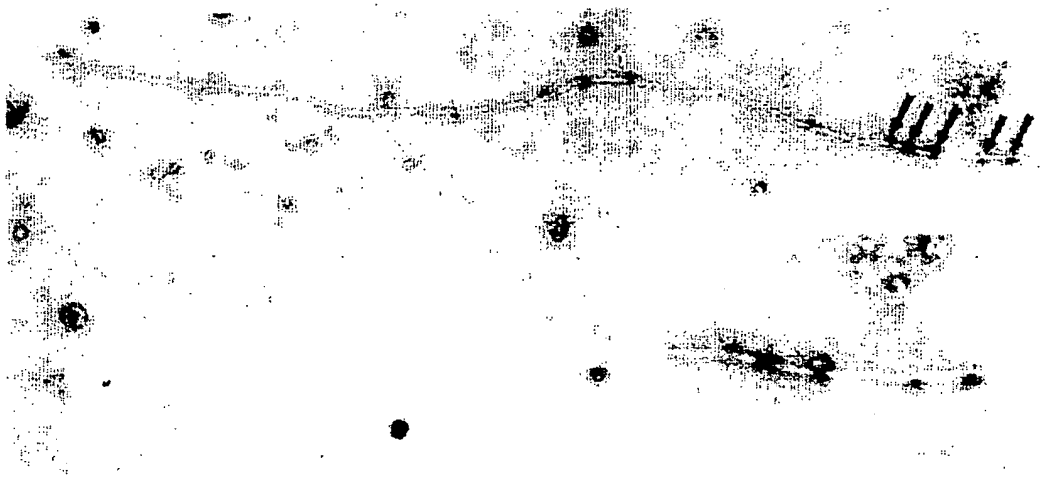


FIG. 9B-1

FIG. 9B-2

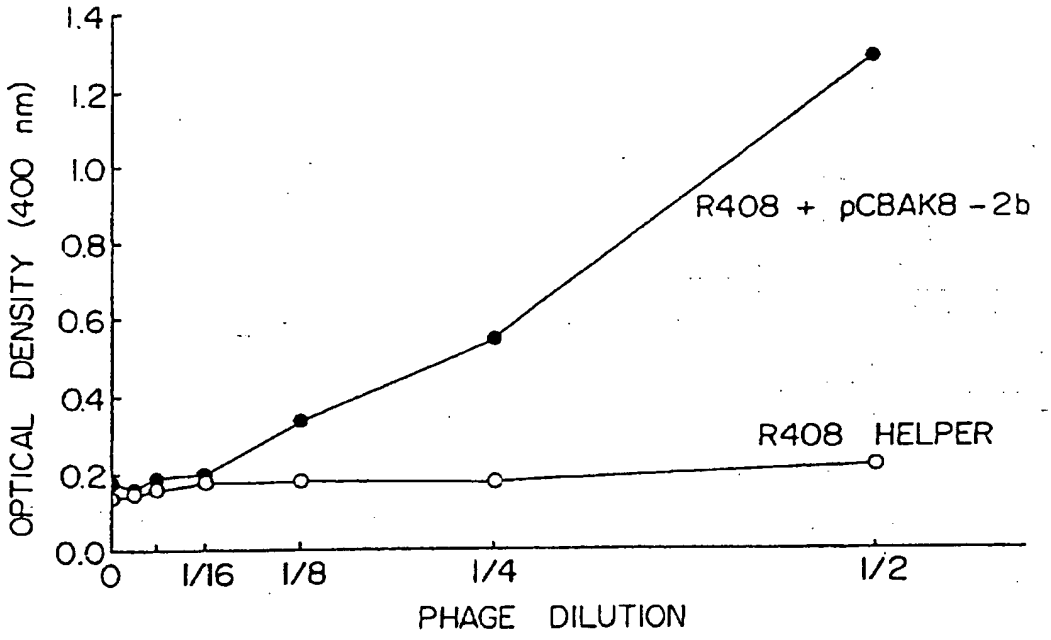

OPTICAL DENSITY (400 nm)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
R408 + pCBAK8 - 2b
R408 HELPER
0
1/16
1/8
1/4
1/2
PHAGE DILUTION

FIG. 10

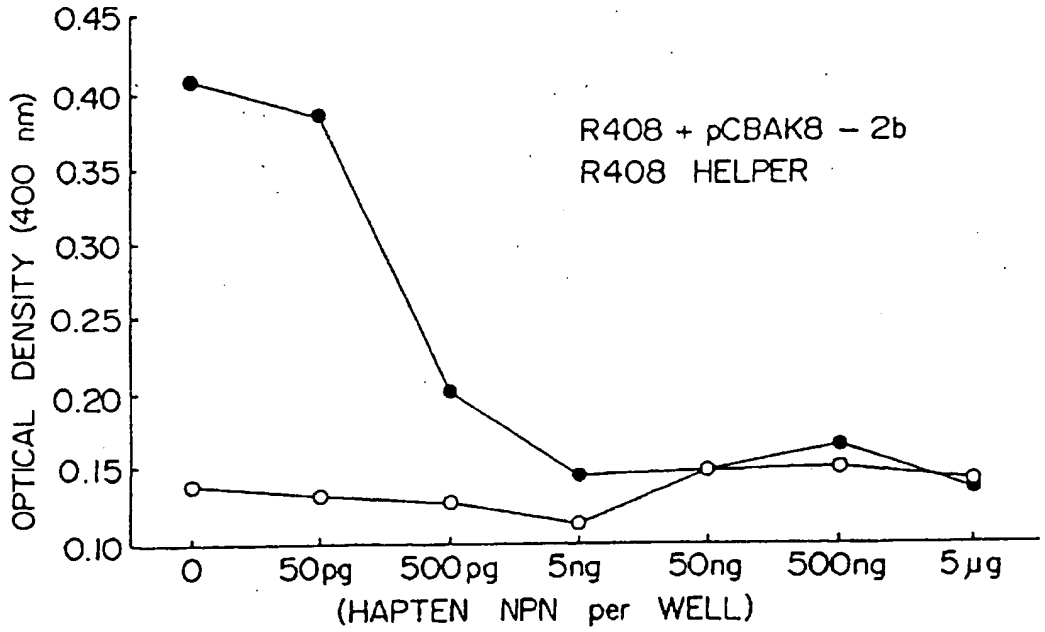

OPTICAL DENSITY (400 nm)
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
R408 + pCBAK8 - 2b
R408 HELPER
0
50pg
500pg
5ng
50ng
500ng
5µg
(HAPTEN NPN per WELL)

FIG. 11

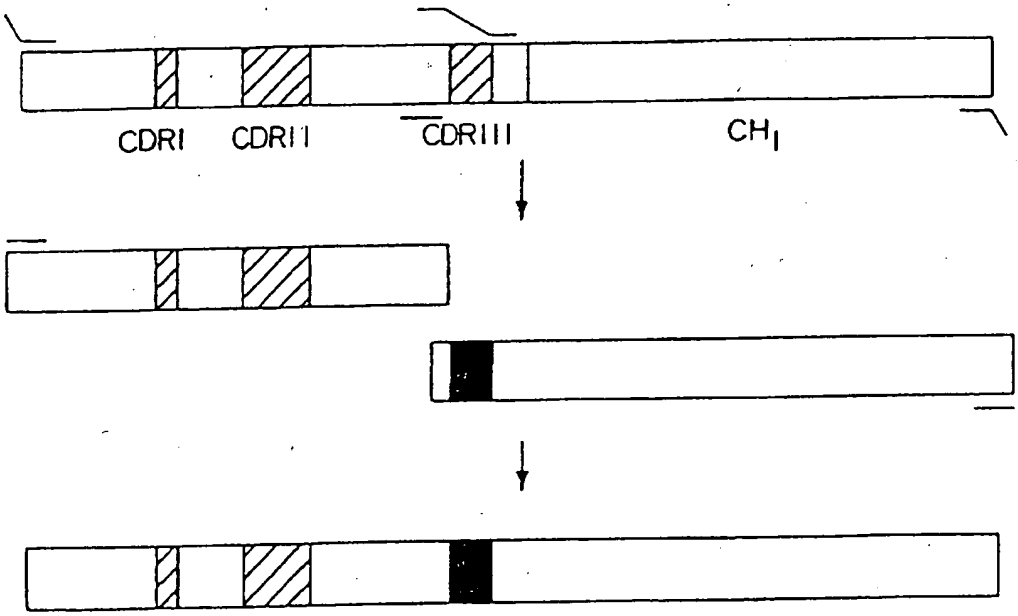


FIG. 12

HETERODIMERIC RECEPTOR LIBRARIES USING PHAGEMIDS

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation of U.S. patent application Ser. No. 07/826,623, filed on Jan. 27, 1992, now abandoned, which is a continuation-in-part of U.S. patent application Ser. No. 07/683,602, filed Apr. 10, 1991, now abandoned, the disclosures of which are hereby incorporated by reference.

This invention was made with government support under Grant No. CA 27489 awarded by the National Institutes of Health.

TECHNICAL FIELD

The present invention relates to cloning vectors and methods for producing a library of DNA molecules capable of expressing a fusion protein on the surface of a filamentous phage particle.

BACKGROUND

Filamentous bacteriophages are a group of related viruses that infect bacteria. They are termed filamentous because they are long and thin particles comprised of an elongated capsule that envelopes the deoxyribonucleic acid (DNA) that forms the bacteriophage genome. The F pili filamentous bacteriophage (Ff phage) infect only gram-negative bacteria by specifically adsorbing to the tip of F pili, and include fd, f1 and M13.

The mature capsule of Ff phage is comprised of a coat of five phage-encoded gene products: cpVIII, the major coat protein product of gene VIII that forms the bulk of the capsule; and four minor coat proteins, cpIII and cpIV at one end of the capsule and cpVII and cpIX at the other end of the capsule. The length of the capsule is formed by 2500 to 3000 copies of cpVIII in an ordered helix array that forms the characteristic filament structure. About five copies each of the minor coat proteins are present at the ends of the capsule. The gene III-encoded protein (cpIII) is typically present in 4 to 6 copies at one end of the capsule and serves as the receptor for binding of the phage to its bacterial host in the initial phase of infection. For detailed reviews of Ff phage structure, see Rasched et al., *Microbiol. Rev.*, 50:401–427 (1986); and Model et al., in "The Bacteriophages, Volume 2", R. Calendar, Ed., Plenum Press, pp. 375–456 (1988).

The assembly of a Ff phage particle involves highly complex mechanics. No phage particles are assembled within a host cell; rather, they are assembled during extrusion of the viral genome through the host cell's membrane. Prior to extrusion, the major coat protein cpVIII and the minor coat protein cpIII are synthesized and transported to the host cell's membrane. Both cpVIII and cpIII are anchored in the host cell membrane prior to their incorporation into the mature particle. In addition, the viral genome is produced and coated with cpV protein. During the extrusion process, cpV-coated genomic DNA is stripped of the cpV coat and simultaneously recoated with the mature coat proteins. The assembly mechanisms that control transferral of these proteins from the membrane to the particle is not presently known.

Both cpIII and cpVIII proteins include two domains that provide signals for assembly of the mature phage particle. The first domain is a secretion signal that directs the newly synthesized protein to the host cell membrane. The secretion signal is located at the amino terminus of the protein and targets the protein at least to the cell membrane. The second domain is a membrane anchor domain that provides signals for association with the host cell membrane and for association with the phage particle during assembly. This second signal for both cpVIII and cpIII comprises at least a hydrophobic region for spanning the membrane.

cpVIII has been extensively studied as a model membrane protein because it can integrate into lipid bilayers such as the cell membrane in an asymmetric orientation with the acidic amino terminus toward the outside and the basic carboxy terminus toward the inside of the membrane. The mature protein is about 50 amino acid residues in length of which 11 residues provide the carboxy terminus, 19 residues provide the hydrophobic transmembrane region, and the remaining residues comprise the amino terminus. Considerable research has been done on the secretion signal region of cpVIII to advance the study of membrane protein synthesis and targeting to membranes. However, little is known about the changes that are tolerated in the structure of the cpVIII membrane anchor region that would allow for assembly of phage particles.

Manipulation of the sequence of cpIII shows that the C-terminal 23 amino acid residue stretch of hydrophobic amino acids normally responsible for a membrane anchor function can be altered in a variety of ways and retain the capacity to associate with membranes. However, those anchor-modified cpIII proteins lost their ability to genetically complement gene III mutants indicating that the requirements of a membrane anchor for functional assembly have not been elucidated.

Ff phage-based expression vectors have been described in which the entire cpIII amino acid residue sequence was modified by insertion of short polypeptide "epitopes" [Parmely et al., *Gene*, 73:305–318 (1988); and Cwirla et al., *Proc. Natl. Acad. Sci. USA*, 87:6378–6382 (1990)] or an amino acid residue sequence defining a single chain antibody domain. McCafferty et al., *Science*, 348:552–554 (1990). These hybrid proteins were synthesized and assembled onto phage particles in amounts of about 5 copies per particle, a density at which normal cpIII is usually found. However, these expressed fusion proteins include the entire cpIII amino acid residue sequence and do not suggest fusion proteins that utilize only the carboxy terminal membrane anchor domain of cpIII.

In addition, no expression system has been described in which a phage coat protein has been engineered to allow assembly of a heteromeric molecule that is functional and capable of incorporation into the coat of a phage particle.

BRIEF SUMMARY OF THE INVENTION

A new, high density, surface-integration technology has been discovered for expressing a recombinant gene product on the surface of a filamentous phage containing the recombinant gene. The invention uses a filamentous phage cpVIII membrane anchor domain as a means for linking gene-product and gene during the assembly stage of filamentous phage replication.

That is, during filamentous phage replication, gene VIII-encoded proteins assemble into a matrix which encapsulates the phage genome. It has now been discovered that (1) phage assembly is not disrupted when recombinant gene VIII-encoded proteins are present, (2) recombinant gene VIII-encoded proteins can be integrated into the assembling matrix, and (3) integration into the matrix can be directed to occur in a surface-accessible orientation.

The present invention can be advantageously applied to the production of heteromeric receptors of predetermined specificity, i.e., it can be used to produce antibodies, T-cell receptors and the like that bind a preselected ligand.

Thus, the present invention provides for linking the functions of heteromeric receptor recognition and filamentous phage replication in a method for isolating a heteromeric receptor. The method produces a filamentous phage comprised of a matrix of gene VIII-encoded proteins that encapsulate a recombinant genome. The recombinant genome contains genes encoding the proteins of the heteromeric receptor proteins. The heteromeric receptor is surface-integrated into the encapsulating matrix via a gene VIII-encoded membrane anchor domain that is fused by a peptide bond during translation to one of the heteromeric receptor proteins. The heteromeric receptor and the genes which encode it are physically linked during the assembly stage of the phage replication cycle. Specifically binding the receptor-coated phage to a solid-support advantageously provides a means for isolating a recombinant genome that encodes a desired heteromeric receptor from a diverse library of recombinant genomes.

In one embodiment, the present invention contemplates an antibody molecule comprising heavy- and light-chain proteins, said heavy-chain protein comprising a V_H-domain flanked by an amino-terminal prokaryotic secretion signal domain and a carboxy-terminal filamentous phage cpVIII membrane anchor domain, said light chain protein comprising a V_L-domain fused to an amino-terminal prokaryotic secretion signal domain.

In another embodiment, the present invention contemplates a vector for expressing a fusion protein, said vector comprising a cassette that includes upstream and downstream translatable DNA sequences operatively linked via a sequence of nucleotides adapted for directional ligation of an insert DNA, said upstream sequence encoding a prokaryotic secretion signal, said downstream sequence encoding a filamentous phage gene cpVIII membrane anchor, said translatable DNA sequences operatively linked to a set of DNA expression signals for expression of said translatable DNA sequences as portions of said fusion protein.

BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings forming a portion of this disclosure:

FIG. 1 illustrates a schematic diagram of the immunoglobulin molecule showing the principal structural features. The circled area on the heavy chain represents the variable region (V_H), a polypeptide containing a biologically active (ligand binding) portion of that region, and a gene coding for that polypeptide, are produced by the methods of the present invention.

FIG. 2A is a diagrammatic sketch of a heavy (H) chain of human IgG (IgG1 subclass). Numbering is from the N-terminus on the left to the C-terminus on the right. Note the presence of four domains, each containing an intrachain disulfide bond (S—S) spanning approximately 60 amino acid residues. The symbol CHO stands for carbohydrate. The V region of the heavy (H) chain (V_H) resembles V_L in having three hypervariable CDR (not shown).

FIG. 2B-1 is a diagrammatic sketch of a human light (Kappa) chain (Panel 1). Numbering is from the N-terminus on the left to the C-terminus on the right. Note the intrachain disulfide bond (S—S) spanning about the same number of amino acid residues in the V_L and C_L domains.

FIG. 2B-2 shows the locations of the complementarity-determining regions (CDR) in the V_L domain. Segments outside the CDR are the framework segments (FR).

FIG. 3 illustrates the sequence of the double-stranded synthetic DNA inserted into Lambda Zap to produce a Lambda Hc2 expression vector. The preparation of the double-stranded synthetic DNA insert is described in Example 1a(ii). The various features required for this vector to express the V_H-coding DNA homologs include the Shine-Dalgarno ribosome binding site, a leader sequence to direct the expressed protein to the periplasm as described by Mouva et al., *J. Biol. Chem.*, 255:27, 1980, and various restriction enzyme sites used to operatively link the V_H homologs to the expression vector. The V_H expression vector sequence also contains a short nucleic acid sequence that codes for amino acids typically found in variable regions heavy chain (V_H Backbone). This V_H Backbone is just upstream and in the proper reading as the V_H DNA homologs that are operatively linked into the Xho I and Spe I cloning sites. The sequences of the top and bottom strands of the double-stranded synthetic DNA insert are listed respectively as SEQ. ID. NO. 1 and SEQ. ID. NO. 2. The synthetic DNA insert is directionally ligated into Lambda Zap II digested with the restriction enzymes Not 1 and Xho I to form Lambda Hc2 expression vector.

FIG. 4 illustrates the major features of the bacterial expression vector Lambda Hc2 (V_H expression vector). The synthetic DNA sequence from FIG. 3 is shown at the top along with the T_3 polymerase promoter from Lambda Zap II. The orientation of the insert in Lambda Zap II is shown. The V_H DNA homologs are inserted into the Xho I and Spe I cloning sites. The read through transcription produces the decapeptide epitope (tag) that is located just 3' of the cloning site.

FIG. 5 illustrates the sequence of the double-stranded synthetic DNA inserted into Lambda Zap to produce a Lambda Lc2 expression vector. The various features required for this vector to express the V_L-coding DNA homologs are described in FIG. 3. The V_L-coding DNA homologs are operatively linked into the Lc2 sequence at the Sac I and Xho I restriction sites. The sequences of the top and bottom strands of the double-stranded synthetic DNA insert are listed respectively as SEQ. ID. NO. 3 and SEQ. ID. NO. 4. The synthetic DNA insert is directionally ligated into Lambda Zap II digested with the restriction enzymes Sac I and Not I to form Lambda Lc2 expression vector.

FIG. 6 illustrates the major features of the bacterial expression vector Lc2 (V_L expression vector). The synthetic DNA sequence from FIG. 5 is shown at the top along with the T_3 polymerase promoter from Lambda Zap II. The orientation of the insert in Lambda Zap II is shown. The V_L DNA homologs are inserted into the Sac I and Xho I cloning sites.

FIG. 7 illustrates the dicistronic expression vector, pComb, in the form of a phagemid expression vector. To produce pComb, phagemids were first excised from the expression vectors, Lambda Hc2 and Lambda Lc2, using an in vivo excision protocol according to manufacturers instructions (Stratagene, La Jolla, Calif.). The pComb expression vector is prepared from Lambda Hc2 and Lambda Lc2 which do not contain V_H-coding or V_L-coding DNA homologs. The in vivo excision protocol moved the cloned insert from the Lambda Hc2 and Lc2 vectors into a phagemid vector. The resultant phagemids contained the same nucleotide sequences for antibody fragment cloning and expression as did the parent vectors. Hc2 and Lc2 phagemid expression vectors were separately restriction digested with Sca I and EcoR I. The linearized phagemids were ligated via the Sca I and EcoR I cohesive termini to form the dicistronic (combinatorial) vector, pComb.

FIG. 8 illustrates a schematic diagram of the composition of pCBAK8-2b phagemid vector, the pathway for Fab assembly and incorporation in phage coat. The vector carries the chloramphenicol acetyl transferase (CAT) marker gene in addition to the nucleotide residue sequences encoding the Fd-cpVIII fusion protein and the kappa chain. The f1 phage origin of replication facilitates the generation of single stranded phagemid. The isopropyl thiogalactopyranoside (IPTG) induced expression of a dicistronic message encoding the Fd-cpVIII fusion (V_H, C_{H1}, cpVIII) and the light chain (V_L, C_L) leads to the formation of heavy and light chains. Each chain is delivered to the periplasmic space by the pelB target sequence, which is subsequently cleaved. The heavy chain is anchored in the membrane by cpVIII fusion while the light chain is secreted into the periplasm. The heavy chain in the presence of light chain assembles to form Fab molecules. The Fabs are incorporated into phage particles via cpVIII (black dots).

FIG. 9 illustrates the electron micrographic localization of 5–7 nm colloidal gold particles coated with NPN-BSA conjugate along the surface of filamentous phage, and from phage emerging from a bacterial cell. Panel 9A shows filamentous phage emerging from the surface of the bacterial cell specifically labelled with the colloidal gold particles coated with BSA-NPN antigen. Panel 9B shows a portion of a mature filamentous phage on the length of which is exhibited the labelling of antigen binding sites.

FIG. 10 illustrates the results of a two-site ELISA for assaying for the presence and function of Fab antibody attached to the surface of bacteriophage particles as described in Example 4b. For expression of Fab antibody on phage surfaces, XL1-Blue cells were transformed with the phagemid expression vector, pCBAK8-2b. The inducer, isopropyl thiogalactopyranoside (IPTG), was admixed with the bacterial suspension at a final concentration of 1 mM for one hour. Helper phage was then admixed with the bacterial suspension to initiate the generation of copies of the sense strand of the phagemid DNA. After a two hour maintenance period, bacterial supernatants containing bacteriophage particles were collected for assaying in ELISA.

Specific titratable binding of NPN-Fab-expressing bacteriophage particles to NPN-coated plates was exhibited. No binding was detected with helper phage alone.

FIG. 11 illustrates the inhibition of NPN-Fab expressing bacteriophage to NPN antigen-coated plates with the addition of increasing amounts of free hapten. The assays were performed as described in FIG. 10. Complete inhibition of binding was observed with 5 ng of added free NPN hapten.

FIG. 12 illustrates schematically the process of mutagenizing the CDR3 region of a heavy chain fragment resulting in an alteration of binding specificity. The oligonucleotide primers are indicated by black bars. The process is described in Example 6.

DETAILED DESCRIPTION OF THE INVENTION

A. Definitions

Amino Acid Residue: An amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are preferably in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH_2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature (described in *J. Biol. Chem.*, 243:3552–59 (1969) and adopted at 37 C.F.R. 1.822(b)(2)). abbreviations for amino acid residues are shown in the following Table of Correspondence:

TABLE OF CORRESPONDENCE

SYMBOL		
1-Letter	3-Letter	AMINO ACID
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine
A	Ala	alanine
S	Ser	serine
I	Ile	isoleucine
L	Leu	leucine
T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
B	Asx	Asn and/or Asp
C	Cys	cysteine
J	Xaa	Unknown or other

It should be noted that all amino acid residue sequences represented herein by formulae have a left-to-right orientation in the conventional direction of amino terminus to carboxy terminus. In addition, the phrase "amino acid residue" is broadly defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those listed in 37 C.F.R. 1.822(b)(4), and incorporated herein by reference. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino acid residues or a covalent bond to an amino-terminal group such as NH_2 or acetyl or to a carboxy-terminal group such as COOH.

Nucleotide: A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleoside. When the nucleoside contains a phosphate group bonded to the 3' or 5' position of the pentose it is referred to as a nucleotide. A sequence of operatively linked nucleotides is typically referred to herein as a "base sequence" or "nucleotide sequence", and their grammatical equivalents, and is represented herein by a formula whose left to right orientation is in the conventional direction of 5'-terminus to 3'-terminus.

Base Pair (bp): A partnership of adenine (A) with thymine (T), or of cytosine (C) with guanine (G) in a double stranded DNA molecule. In RNA, uracil (U) is substituted for thymine.

Nucleic Acid: A polymer of nucleotides, either single or double stranded.

Polynucleotide: a polymer of single or double stranded nucleotides. As used herein "polynucleotide" and its grammatical equivalents will include the full range of nucleic acids. A polynucleotide will typically refer to a nucleic acid

molecule comprised of a linear strand of two or more deoxyribonucleotides and/or ribonucleotides. The exact size will depend on many factors, which in turn depends on the ultimate conditions of use, as is well known in the art. The polynucleotides of the present invention include primers, probes, RNA/DNA segments, oligonucleotides or "oligos" (relatively short polynucleotides), genes, vectors, plasmids, and the like.

Gene: A nucleic acid whose nucleotide sequence codes for an RNA or polypeptide. A gene can be either RNA or DNA.

Duplex DNA: a double-stranded nucleic acid molecule comprising two strands of substantially complementary polynucleotides held together by one or more hydrogen bonds between each of the complementary bases present in a base pair of the duplex. Because the nucleotides that form a base pair can be either a ribonucleotide base or a deoxyribonucleotide base, the phrase "duplex DNA" refers to either a DNA-DNA duplex comprising two DNA strands (ds DNA), or an RNA-DNA duplex comprising one DNA and one RNA strand.

Complementary Bases: Nucleotides that normally pair up when DNA or RNA adopts a double stranded configuration.

Complementary Nucleotide Sequence: A sequence of nucleotides in a single-stranded molecule of DNA or RNA that is sufficiently complementary to that on another single strand to specifically hybridize to it with consequent hydrogen bonding.

Conserved: A nucleotide sequence is conserved with respect to a preselected (reference) sequence if it non-randomly hybridizes to an exact complement of the preselected sequence.

Hybridization: The pairing of substantially complementary nucleotide sequences (strands of nucleic acid) to form a duplex or heteroduplex by the establishment of hydrogen bonds between complementary base pairs. It is a specific, i.e. non-random, interaction between two complementary polynucleotides that can be competitively inhibited.

Nucleotide Analog: A purine or pyrimidine nucleotide that differs structurally from A, T, G, C, or U, but is sufficiently similar to substitute for the normal nucleotide in a nucleic acid molecule.

DNA Homolog: Is a nucleic acid having a preselected conserved nucleotide sequence and a sequence coding for a receptor capable of binding a preselected ligand.

Recombinant DNA (rDNA) molecule: a DNA molecule produced by operatively linking two DNA segments. Thus, a recombinant DNA molecule is a hybrid DNA molecule comprising at least two nucleotide sequences not normally found together in nature. rDNA's not having a common biological origin, i.e., evolutionarily different, are said to be "heterologous".

Vector: a rDNA molecule capable of autonomous replication in a cell and to which a DNA segment, e.g., gene or polynucleotide, can be operatively linked so as to bring about replication of the attached segment. Vectors capable of directing the expression of genes encoding for one or more proteins are referred to herein as "expression vectors". Particularly important vectors allow cloning of cDNA (complementary DNA) from mRNAs produced using reverse transcriptase.

Receptor: A receptor is a molecule, such as a protein, glycoprotein and the like, that can specifically (non-randomly) bind to another molecule.

Antibody: The term antibody in its various grammatical forms is used herein to refer to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antibody combining site or paratope. Exemplary antibody molecules are intact immunoglobulin molecules, substantially intact immunoglobulin molecules and portions of an immunoglobulin molecule, including those portions known in the art as Fab, Fab', $F(ab')_2$ and F(v).

Antibody Combining Site: An antibody combining site is that structural portion of an antibody molecule comprised of a heavy and light chain variable and hypervariable regions that specifically binds (immunoreacts with) an antigen. The term immunoreact in its various forms means specific binding between an antigenic determinant-containing molecule and a molecule containing an antibody combining site such as a whole antibody molecule or a portion thereof.

Monoclonal Antibody: The phrase monoclonal antibody in its various grammatical forms refers to a population of antibody molecules that contains only one species of antibody combining site capable of immunoreacting with a particular antigen. A monoclonal antibody thus typically displays a single binding affinity for any antigen with which it immunoreacts. A monoclonal antibody may therefore contain an antibody molecule having a plurality of antibody combining sites, each immunospecific for a different antigen, e.g., a bispecific monoclonal antibody.

Fusion Protein: A protein comprised of at least two polypeptides and a linking sequence to operatively link the two polypeptides into one continuous polypeptide. The two polypeptides linked in a fusion protein are typically derived from two independent sources, and therefore a fusion protein comprises two linked polypeptides not normally found linked in nature.

Upstream: In the direction opposite to the direction of DNA transcription, and therefore going from 5' to 3' on the non-coding strand, or 3' to 5' on the mRNA.

Downstream: Further along a DNA sequence in the direction of sequence transcription or read out, that is traveling in a 3'- to 5'-direction along the non-coding strand of the DNA or 5'- to 3'-direction along the RNA transcript.

Cistron: Sequence of nucleotides in a DNA molecule coding for an amino acid residue sequence and including upstream and downstream DNA expression control elements.

Stop Codon: Any of three codons that do not code for an amino acid, but instead cause termination of protein synthesis. They are UAG, UAA and UGA and are also referred to as a nonsense or termination codon.

Leader Polypeptide: A short length of amino acid sequence at the amino end of a protein, which carries or directs the protein through the inner membrane and so ensures its eventual secretion into the periplasmic space and perhaps beyond. The leader sequence peptide is commonly removed before the protein becomes active.

Reading Frame: Particular sequence of contiguous nucleotide triplets (codons) employed in translation. The reading frame depends on the location of the translation initiation codon.

B. Filamentous Phage

The present invention contemplates a filamentous phage comprising a matrix of cpVIII proteins encapsulating a genome encoding first and second polypeptides. The phage further comprises a heteromeric receptor comprised of the first and second polypeptides surface-integrated into the matrix via a cpVIII membrane anchor domain fused to at

least one of the first or second polypeptides. Preferably, the first and second polypeptides are V_H and V_L proteins, respectively.

The first and second polypeptides are capable of autogenous assembly into a functional receptor, which is expressed on the outer surface in a manner accessible to ligand, i.e. they are surface-integrated into the phage. Typically, the receptor is comprised of a linking polypeptide that contains the cpVIII membrane anchor domain, such as a polypeptide described in Section C, and a non-linking polypeptide(s).

Because the receptor is linked to the phage in a surface accessible manner, the phage can be advantageously used as a solid-phase affinity sorbent. In preferred embodiments, the phage are linked, preferably removably linked, to a solid (aqueous insoluble) matrix such as agarose, cellulose, synthetic reins, polysaccharides and the like. For example, transformants shedding the phage can be applied to and retained in a column and maintained under conditions that support shedding of the phage. An aqueous composition containing a ligand that binds to the receptor expressed by the phage is then passed through the column at a predetermined rate and under receptor-binding conditions to form a solid-phase receptor-ligand complex. The column is then washed to remove unbound material, leaving the ligand bound to the solid-phase phage. The ligand can then be removed and recovered by washing the column with a buffer that promotes dissociation of the receptor-ligand complex.

Alternatively, purified phage can be admixed with a aqueous solution containing the ligand to be affinity purified. The receptor/ligand binding reaction admixture thus formed is maintained for a time period and under binding conditions sufficient for a phage-linked receptor-ligand complex to form. The phage-bound ligand (ligand-bearing phage) are then separated and recovered from the unbound materials, such as by centrifugation, electrophoresis, precipitation, and the like.

C. DNA Expression Vectors

A vector of the present invention is a recombinant DNA (rDNA) molecule adapted for receiving and expressing translatable DNA sequences in the form of a fusion protein containing a filamentous phage gene VIII membrane anchor domain and a prokaryotic secretion signal domain. The vector comprises a cassette that includes upstream and downstream translatable DNA sequences operatively linked via a sequence of nucleotides adapted for directional ligation. The upstream translatable sequence encodes the secretion signal. The downstream translatable sequence encodes the filamentous phage membrane anchor. The cassette preferably includes DNA expression control sequences for expressing the fusion protein that is produced when a translatable DNA sequence is directionally inserted into the cassette via the sequence of nucleotides adapted for directional ligation.

An expression vector is characterized as being capable of expressing, in a compatible host, a structural gene product such as a fusion protein of the present invention.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting between different genetic environments another nucleic acid to which it has been operatively linked. Preferred vectors are those capable of autonomous replication and expression of structural gene products present in the DNA segments to which they are operatively linked.

As used herein with regard to DNA sequences or segments, the phrase "operatively linked" means the sequences or segments have been covalently joined into one shard of DNA, whether in single or double stranded form.

The choice of vector to which a cassette of this invention is operatively linked depends directly, as is well known in the art, on the functional properties desired, e.g., vector replication and protein expression, and the host cell to be transformed, these being limitations inherent in the art of constructing recombinant DNA molecules.

In preferred embodiments, the vector utilized includes a prokaryotic replicon i.e., a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extra chromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, those embodiments that include a prokaryotic replicon also include a gene whose expression confers a selective advantage, such as drug resistance, to a bacterial host transformed therewith. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline. Vectors typically also contain convenient restriction sites for insertion of translatable DNA sequences. Exemplary vectors are the plasmids pUC8, pUC9, pBR322, and pBR329 available from BioRad Laboratories, (Richmond, Calif.) and pPL and pKK223 available from Pharmacia, (Piscataway, N.J.).

A sequence of nucleotides adapted for directional ligation, i.e., a polylinker, is a region of the DNA expression vector that (1) operatively links for replication and transport the upstream and downstream translatable DNA sequences and (2) provides a site or means for directional ligation of a DNA sequence into the vector. Typically, a directional polylinker is a sequence of nucleotides that defines two or more restriction endonuclease recognition sequences, or restriction sites. Upon restriction cleavage, the two sites yield cohesive termini to which a translatable DNA sequence can be ligated to the DNA expression vector. Preferably, the two restriction sites provide, upon restriction cleavage, cohesive termini that are non-complementary and thereby permit directional insertion of a translatable DNA sequence into the cassette. In one embodiment, the directional ligation means is provided by nucleotides present in the upstream translatable DNA sequence, downstream translatable DNA sequence, or both. In another embodiment, the sequence of nucleotides adapted for directional ligation comprises a sequence of nucleotides that defines multiple directional cloning means. Where the sequence of nucleotides adapted for directional ligation defines numerous restriction sites, it is referred to as a multiple cloning site.

A translatable DNA sequence is a linear series of nucleotides that provide an uninterrupted series of at least 8 codons that encode a polypeptide in one reading frame.

An upstream translatable DNA sequence encodes a prokaryotic secretion signal. The secretion signal is a leader peptide domain of protein that targets the protein to the periplasmic membrane of gram negative bacteria.

A preferred secretion signal is a pelB secretion signal. The predicted amino acid residue sequences of the secretion signal domain from two pelB gene product variants from *Erwinia carotova* are shown in Table 1 as described by Lei, et al., *Nature*, 331:543–546 (1988). A particularly preferred pelB secretion signal is also shown in Table 1.

The leader sequence of the pelB protein has previously been used as a secretion signal for fusion proteins. Better et al., *Science*, 240:1041–1043 (1988); Sastry et al., *Proc. Natl. Acad. Sci. USA*, 86:5728–5732 (1989); and Mullinax et al., *Proc. Natl. Acad. Sci. USA*, 87:8095–8099 (1990).

Amino acid residue sequences for other secretion signal polypeptide domains from *E. coli* useful in this invention are also listed in Table 1. Oliver. In Neidhard. F. C. (ed.). *Escherichia coli* and *Salmonella Typhimurium*. American Society for Microbiology. Washington. D.C.. 1:56–69 (1987).

A translatable DNA sequence encoding the pelB secretion signal having the amino acid residue sequence shown in SEQ. ID. NO. 5 is a preferred DNA sequence for inclusion in a DNA expression vector of

TABLE 1

Leader Sequences

SEQ. ID. NO.	Type	Amino Acid Residue Sequence
(5)	PelB[1]	MetLysTyrLeuLeuProThrAlaAlaAlaGlyLeuLeu LeuLeuAlaAlaGlnProAlaMet
(6)	PelB[2]	MetLysTyrLeuLeuProThrAlaAlaAlaGlyLeuLeu LeuLeuAlaAlaGlnProAlaGlnProAlaMetAla
(7)	PelB[3]	MetLysSerLeuIleThrProIleAlaAlaGlyLeuLeu LeuAlaPheSerGlnTyrSerLeuAla
(8)	MalE[4]	MetLysIleLysThrGlyAlaArgIleLeuAlaLeuSer AlaLeuThrThrMetMetPheSerAlaSerAlaLeuAla LysIle
(9)	OmpF[4]	MetMetLysArgAsnIleLeuAlaValIleValProAla LeuLeuValAlaGlyThrAlaAsnAlaAlaGlu
(10)	PhoA[4]	MetLysGlnSerThrIleAlaLeuAlaLeuLeuProLeu LeuPheThrProValThrLysAlaArgThr
(11)	Bla[4]	MetSerIleGlnHisPheArgValAlaLeuIleProPhe PheAlaAlaPheCysLeuProValPheAlaHisPro
(12)	LamB[4]	MetMetIleThrLeuArgLysLeuProLeuAlaValAla ValAlaAlaGlyValMetSerAlaGlnAlaMetAlaVal Asp
(13)	Lpp[4]	MetLysAlaThrLysLeuValLeuGlyAlaValIleLeu GlySerThrLeuLeuAlaGlyCysSer
(14)	cpVIII[5]	MetLysLysSerLeuValLeuLysAlaSerValAlaVal AlaThrLeuValProMetLeuSerPheAla
(15)	cpIII[6]	MetLysLysLeuLeuPheAlaIleProLeuValValPro PheTyrSerHisSer

[1]pelB used in this invention
[2]pelB from *Erwinia carotovora* gene
[3]pelB from *Erwinia carotovora* EC 16 gene
[4]leader sequences from *E. coli*
[5]leader sequence for cpVIII
[6]leader sequence for cpIII

this invention.

A downstream translatable DNA sequence encodes a filamentous phage membrane anchor. Preferred membrane anchors are obtainable from filamentous phage M13, f1, fd, and the like. Preferred membrane anchor domains are found in the coat proteins encoded by gene III and gene VIII. Thus, a downstream translatable DNA sequence encodes an amino acid residue sequence that corresponds, and preferably is identical, to the membrane anchor domain of either a filamentous phage gene III or gene VIII coat protein.

The membrane anchor domain of a filamentous phage coat protein is the carboxy terminal region of the coat protein and includes a region of hydrophobic amino acid residues for spanning a lipid bilayer membrane, and a region of charged amino acid residues normally found at the cytoplasmic face of the membrane and extending away from the membrane. In the phage f1, gene VIII coat protein's membrane spanning region comprises residue Trp-26 through Lys-40, and the cytoplasmic region comprises the carboxy-terminal 11 residues from 41 to 52. Ohkawa et al., *J. Biol. Chem.*, 256:9951–9958 (1981).

The amino acid residue sequence of a preferred membrane anchor domain derived from the M13 filamentous phage gene III coat protein (also designated cpIII) has a sequence shown in SEQ. ID. NO. 16 from residue 1 to residue 211. Gene III coat protein is present on a mature filamentous phage at one end of the phage particle with typically about 4 to 6 copies of the coat protein.

The amino acid residue sequence of a preferred membrane anchor domain derived from the M13 filamentous phage gene VIII coat protein (also designated cpVIII) has a sequence shown in SEQ. ID. NO. 17 from residue 1 to residue 50. Gene VIII coat protein is present on a mature filamentous phage over the majority of the phage particle with typically about 2500 to 3000 copies of the coat protein.

For detailed descriptions of the structure of filamentous phage particles, their coat proteins and particle assembly, see the reviews by Rached et al., *Microbiol. Rev.*, 50:401–427 (1986); and Model et al., in "The Bacteriophages: Vol. 2", R. Calendar, ed. Plenum Publishing Co., pp. 375–456, (1988).

A cassette in a DNA expression vector of this invention is the region of the vector that forms, upon insertion of a translatable DNA sequence, a sequence of nucleotides capable of expressing, in an appropriate host, a fusion protein of this invention. The expression-competent sequence of nucleotides is referred to as a cistron. Thus, the cassette comprises DNA expression control elements operatively linked to the upstream and downstream translatable DNA sequences. A cistron is formed when a translatable DNA sequence is directionally inserted (directionally ligated) between the upstream and downstream sequences via the sequence of nucleotides adapted for that purpose. The resulting three translatable DNA sequences, namely the upstream, the inserted and the downstream sequences, are all operatively linked in the same reading frame.

DNA expression control sequences comprise a set of DNA expression signals for expressing a structural gene product and include both 5' and 3' elements, as is well known, operatively linked to the cistron such that the cistron is able to express a structural gene product. The 5' control sequences define a promoter for initiating transcription and a ribosome binding site operatively linked at the 5' terminus of the upstream translatable DNA sequence.

To achieve high levels of gene expression in *E. coli*, it is necessary to use not only strong promoters to generate large quantities of mRNA, but also ribosome binding sites to ensure that the mRNA is efficiently translated. In *E. coli*, the ribosome binding site includes an initiation codon (AUG) and a sequence 3–9 nucleotides long located 3–11 nucleotides upstream from the initiation codon [Shine et al., *Nature*, 254:34 (1975)] The sequence, AGGAGGU, which is called the Shine-Dalgarno (SD) sequence, is complementary to the 3' end of *E. coli* 16S mRNA. Binding of the ribosome to mRNA and the sequence at the 3' end of the mRNA can be affected by several factors:

(i) The degree of complementarity between the SD sequence and 3' end of the 16S tRNA.

(ii) The spacing and possibly the DNA sequence lying between the SD sequence and the AUG [Roberts et al., *Proc. Natl. Acad. Sci. USA*, 76:760 (1979a); Roberts et al., *Proc. Natl. Acad. Sci. USA*, 76:5596 (1979b); Guarente et al., *Science*, 209:1428 (1980); and Guarente et al., *Cell*, 20:543 (1980).] Optimization is achieved by measuring the level of expression of genes in plasmids in which this spacing is systematically altered. Comparison of different mRNAs shows that there are statistically preferred sequences from positions −20 to +13 (where the A of the AUG is position 0) [Gold et al., *Annu. Rev. Microbiol.*, 35:365 (1981)]. Leader sequences have been shown to influence translation dramatically (Roberts et al., 1979 a, b supra).

(iii) The nucleotide sequence following the AUG, which affects ribosome binding [Taniguchi et al., *J. Mol. Biol.*, 118:533 (1978)].

Useful ribosome binding sites are shown in Table 2 below.

TABLE 2

Ribosome Binding Sites*

SEQ. ID. NO.		
1.	(18)	5' AAUCUUGGAGGCUUUUUUAUGGUUCGUUCU
2.	(19)	5' UAACUAAGGAUGAAAUGCAUGUCUAAGACA
3.	(20)	5' UCCUAGGAGGUUUGACCUAUGCGAGCUUUU
4.	(21)	5' AUGUACUAAGGAGGUUGUAUGGAACAACGC

*Sequences of initiation regions for protein synthesis in four phage mRNA molecules are underlined.
AUG = initiation codon (double underlined)
1. = Phage ϕX174 gene-A protein
2. = Phage Qβ replicase
3. = Phage R17 gene-A protein
4. = Phage lambda gene-cro protein

The 3' control sequences define at least one termination (stop) codon in frame with and operatively linked to the downstream translatable DNA sequence.

Thus, a DNA expression vector of this invention provides a system for cloning translatable DNA sequences into the cassette portion of the vector to produce a cistron capable of expressing a fusion protein of this invention.

In preferred embodiments, a DNA expression vector provides a system for independently cloning two translatable DNA sequences into two separate cassettes present in the vector, to form two separate cistrons for expressing both polypeptides of a heterodimeric receptor, or the ligand binding portions of the polypeptides that comprise a heterodimeric receptor. The DNA expression vector for expressing two cistrons is referred to as a dicistronic expression vector.

Thus, a preferred DNA expression vector of this invention comprises, in addition to the cassette previously described in detail, a second cassette for expressing a second fusion protein. The second cassette includes a third translatable DNA sequence that encodes a secretion signal, as defined herein before, operatively linked at its 3' terminus via a sequence of nucleotides adapted for directional ligation to a downstream DNA sequence defining at least one stop codon. The third translatable DNA sequence is operatively linked at its 5' terminus to DNA expression control sequences forming the 5' elements defined above. The second cassette is capable, upon insertion of a translatable DNA sequence, of expressing the second fusion protein.

In a preferred embodiment, a DNA expression vector is designed for convenient manipulation in the form of a filamentous phage particle according to the teachings of the present invention. In this embodiment, a DNA expression vector further contains a nucleotide sequence that defines a filamentous phage origin of replication such that the vector, upon presentation of the appropriate genetic complementation, can replicate as a filamentous phage in single stranded replicative form and be packaged into filamentous phage particles. This feature provides the ability of the DNA expression vector to be packaged into phage particles for subsequent segregation of the particle, and vector contained therein, away from other particles that comprise a population of phage particles.

A filamentous phage origin of replication is a region of the phage genome, as is well known, that defines sites for initiation of replication, termination of replication and packaging of the replicative form produced by replication. See, for example, Rasched et al., *Microbiol. Rev.*, 50:401–427 (1986); and Horiuchi, *J. Mol. Biol.*, 188:215–223 (1986).

A preferred filamentous phage origin of replication for use in the present invention is a M13, f1 or fd phage origin of replication. Particularly preferred is a filamentous phage origin of replication having a sequence shown in SEQ. ID. NO. 117 and described by Short et al., *Nucl. Acids Res.*, 16:7583–7600 (1988). Preferred DNA expression vectors are the dicistronic expression vectors pCOMB8 and pCOMB3 described in Example 1b(i) and 1b(ii), respectively.

D. Polypeptides

In another embodiment, the present invention contemplates a polypeptide comprising an insert domain flanked by an amino-terminal secretion signal domain and a carboxy-terminal filamentous phage coat protein membrane anchor domain.

Preferably, the polypeptide is a fusion protein having a receptor domain comprised of an amino acid residue sequence that defines the ligand binding domain of a receptor protein positioned between a prokaryotic secretion signal domain and a gene VIII-encoded (cpVIII) membrane anchor domain. In preferred embodiments, the receptor protein is a polypeptide chain of a heterodimeric receptor. Insofar as the polypeptide has a receptor domain, it is also referred to herein as a receptor. In other preferred embodiments the secretion signal domain is a pelB secretion signal as described herein.

Preferred heterodimeric receptors include immunoglobulins, major histocompatibility antigens of class I or II, lymphocyte receptors, integrins and the like heterodimeric receptors.

In one embodiment, a polypeptide of this invention has an amino acid residue sequence that can be represented by the formula, shown in the direction of amino- to carboxy terminus:

$$NH_2—O—(U)_m—V—(X)_n—Z—COOH, \quad (F1)$$

where O represents an amino acid residue sequence defining a secretion signal, U represents a first spacer polypeptide, V represents an amino acid residue sequence defining a receptor domain, X represents a second spacer polypeptide, and Z represents an amino acid residue sequence defining a filamentous phage coat protein membrane anchor, with the proviso that m is the integer 0 or 1 such that when m is O, U is not present, and when m is 1, U is present, and n is 0 or 1 such that when n is O, X is not present and when n is 1, X is present.

In the formula (F1), the secretion signal and the filamentous phage coat protein membrane anchor are as defined herein above. Particularly preferred is a polypeptide according to formula (F1) where Z defines the gene VIII membrane anchor as described herein. In another preferred embodiment the secretion signal is the pelB secretion signal.

In one embodiment, V is an amino acid residue sequence that defines the ligand binding domain of a chain of a heterodimeric receptor molecule, and preferably is an immunoglobulin variable region polypeptide. In a particularly preferred polypeptide V is a V_H or V_L polypeptide. Most preferred is a polypeptide where V is an immunoglobulin V_H polypeptide, and m and n are both zero.

In another embodiment, U or X can define a proteolytic cleavage site, such as the sequence of amino acids found in a precursor protein, such as prothrombin, factor X and the like, that defines the site of cleavage of the protein. A fusion protein having a cleavage site provides a means to purify the protein away from the phage particle to which it is attached.

The polypeptide spacers U and X can each have any sequence of amino acid residues of from about 1 to 6 amino

acid residues in length. Typically the spacer residues are present in a polypeptide to accommodate the continuous reading frame that is required when a polypeptide is produced by the methods disclosed herein using a DNA expression vector of this invention.

A receptor of the present invention assumes a conformation having a binding site specific for, as evidenced by its ability to be competitively inhibited, a preselected or predetermined ligand such as an antigen, enzymatic substrate and the like. In one embodiment, a receptor of this invention is a ligand binding polypeptide that forms an antigen binding site which specifically binds to a preselected antigen to form a complex having a sufficiently strong binding between the antigen and the binding site for the complex to be isolated. When the receptor is an antigen binding polypeptide its affinity or avidity is generally greater than 10^5 M^{-1} more usually greater than 10^6 and preferably greater than 10^8 M^{-1}.

In another embodiment, a receptor of the subject invention binds a substrate and catalyzes the formation of a product from the substrate. While the topology of the ligand binding site of a catalytic receptor is probably more important for its preselected activity than its affinity (association constant or pKa) for the substrate, the subject catalytic receptors have an association constant for the preselected substrate generally greater than 10^3 M^{-1}, more usually greater than 10^5 M^{-1} or 10^6 M^{-1} and preferably greater than 10^7 M^{-1}.

Preferably the receptor produced by the subject invention is heterodimeric and is therefore normally comprised of two different polypeptide chains, which together assume a conformation having a binding affinity, or association constant for the preselected ligand that is different, preferably higher, than the affinity or association constant of either of the polypeptides alone, i.e., as monomers. One or both of the different polypeptide chains is derived from the variable region of the light and heavy chains of an immunoglobulin. Typically, polypeptides comprising the light (V_L) and heavy (V_H) variable regions are employed together for binding the preselected ligand.

A receptor produced by the subject invention can be active in monomeric as well as multimeric forms, either homomeric or heteromeric, preferably heterodimeric. For example, V_H and V_L ligand binding polypeptide produced by the present invention can be advantageously combined in the heterodimer to modulate the activity of either or to produce an activity unique to the heterodimer.

The individual ligand polypeptides will be referred to as V_H and V_L and the heterodimer will be referred to as a Fv. However, it should be understood that a V_H may contain in addition to the V_H, substantially all or a portion of the heavy chain constant region. Similarly, a V_L may contain, in addition to the V_L, substantially all or a portion of the light chain constant region. A heterodimer comprised of a V_H containing a portion of the heavy chain constant region and a V_L containing substantially all of the light chain constant region is termed a Fab fragment. The production of Fab can be advantageous in some situations because the additional constant region sequences contained in a Fab as compared to a Fv can stabilize the V_H and V_L interaction. Such stabilization can cause the Fab to have higher affinity for antigen. In addition the Fab is more commonly used in the art and thus there are more commercial antibodies available to specifically recognize a Fab in screening procedures.

The individual V_H and V_L polypeptides can be produced in lengths equal to or substantially equal to their naturally occurring lengths. However, in preferred embodiments, the V_H and V_L polypeptides will generally have fewer than 125 amino acid residues, more usually fewer than about 120 amino acid residues, while normally having greater than 60 amino acid residues, usually greater than about 95 amino acid residues, more usually greater than about 100 amino acid residues. Preferably, the V_H will be from about 110 to about 230 amino acid residues in length while V_L will be from about 95 to about 214 amino acid residues in length. V_H and V_L chains sufficiently long to form Fabs are preferred.

The amino acid residue sequences will vary widely, depending upon the particular idiotype involved. Usually, there will be at least two cysteines separated by from about 60 to 75 amino acid residues and joined by a disulfide bond. The polypeptides produced by the subject invention will normally be substantial copies of idiotypes of the variable regions of the heavy and/or light chains of immunoglobulins, but in some situations a polypeptide may contain random mutations in amino acid residue sequences in order to advantageously improve the desired activity.

In some situations, it is desirable to provide for covalent cross linking of the V_H and V_L polypeptides, which can be accomplished by providing cysteine resides at the carboxyl termini. The polypeptide will normally be prepared free of the immunoglobulin constant regions, however a small portion of the J region may be included as a result of the advantageous selection of DNA synthesis primers. The D region will normally be included in the transcript of the V_H.

Typically the C terminus region of the V_H and V_L polypeptides will have a greater variety of sequences than the N terminus and, based on the present strategy, can be further modified to permit a variation of the normally occurring V_H and V_L chains. A synthetic polynucleotide can be employed to vary one or more amino acid in a hypervariable region.

In another embodiment, the invention contemplates a heterodimeric receptor molecule comprised of two polypeptide chains, at least one of which is a polypeptide of this invention. Preferably, the polypeptide comprises a receptor domain derived from an immunoglobulin variable chain, more preferably a V_H. More preferred is a heterodimeric receptor comprising receptor domains from both V_H and V_L chains.

E. Methods for Producing a Library

1. General Rationale

In one embodiment the present invention provides a system for the simultaneous cloning and screening of preselected ligand-binding specificities from gene repertoires using a single vector system. This system provides linkage of cloning and screening methodologies and has two requirements. First, that expression of the polypeptide chains of a heterodimeric receptor in an in vitro expression host such as *E. coli* requires coexpression of the two polypeptide chains in order that a functional heterodimeric receptor can assemble to produce a receptor that binds ligand. Second, that screening of isolated members of the library for a preselected ligand-binding capacity requires a means to correlate the binding capacity of an expressed receptor molecule with a convenient means to isolate the gene that encodes the member from the library.

Linkage of expression and screening is accomplished by the combination of targeting of a fusion protein into the periplasm of a bacterial cell to allow assembly of a functional receptor, and the targeting of a fusion protein onto the coat of a filamentous phage particle during phage assembly to allow for convenient screening of the library member of interest. Periplasmic targeting is provided by the presence of a secretion signal domain in a fusion protein of this inven-

tion. Targeting to a phage particle is provided by the presence of a filamentous phage coat protein membrane anchor domain in a fusion protein of this invention.

The present invention describes in one embodiment a method for producing a library of DNA molecules, each DNA molecule comprising a cistron for expressing a fusion protein on the surface of a filamentous phage particle. The method comprises the steps of (a) forming a ligation admixture by combining in a ligation buffer (i) a repertoire of polypeptide encoding genes and (ii) a plurality of DNA expression vectors in linear form adapted to form a fusion protein expressing cistron, and (b) subjecting the admixture to ligation conditions for a time period sufficient for the repertoire of genes to become operatively linked (ligated) to the plurality of vectors to form the library.

In this embodiment, the repertoire of polypeptide encoding genes are in the form of double-stranded (ds) DNA and each member of the repertoire has cohesive termini adapted for directional ligation. In addition, the plurality of DNA expression vectors are each linear DNA molecules having upstream and downstream cohesive termini that are (a) adapted for directionally receiving the polypeptide genes in a common reading frame, and (b) operatively linked to respective upstream and downstream translatable DNA sequences. The upstream translatable DNA sequence encodes a secretion signal, preferably a pelB secretion signal, and the downstream translatable DNA sequence encodes a filamentous phage coat protein membrane anchor as described herein for a polypeptide of this invention. The translatable DNA sequences are also operatively linked to respective upstream and downstream DNA expression control sequences as defined for a DNA expression vector described herein.

The library so produced can be utilized for expression and screening of the fusion proteins encoded by the resulting library of cistrons represented in the library by the expression and screening methods described herein.

2. Production of Gene Repertoires

A gene repertoire is a collection of different genes, preferably polypeptide-encoding genes (polypeptide genes), and may be isolated from natural sources or can be generated artificially. Preferred gene repertoires are comprised of conserved genes. Particularly preferred gene repertoires comprise either or both genes that code for the members of a dimeric receptor molecule.

A gene repertoire useful in practicing the present invention contains at least 10^3, preferably at least 10^4, more preferably at least 10^5, and most preferably at least 10^7 different genes. Methods for evaluating the diversity of a repertoire of genes is well known to one skilled in the art.

Thus, in one embodiment, the present invention contemplates a method of isolating a pair of genes coding for a dimeric receptor having a preselected activity from a repertoire of conserved genes. Additionally, expressing the cloned pair of genes and isolating the resulting expressed dimeric receptor protein is also described. Preferably, the receptor will be a heterodimeric polypeptide capable of binding a ligand, such as an antibody molecule or immunologically active portion thereof, a cellular receptor, or a cellular adhesion protein coded for by one of the members of a family of conserved genes, i.e., genes containing a conserved nucleotide sequence of at least about 10 nucleotides in length.

Exemplary conserved gene families encoding different polypeptide chains of a dimeric receptor are those coding for immunoglobulins, major histocompatibility complex antigens of class I or II, lymphocyte receptors, integrins and the like.

A gene can be identified as belonging to a repertoire of conserved genes using several methods. For example, an isolated gene may be used as a hybridization probe under low stringency conditions to detect other members of the repertoire of conserved genes present in genomic DNA using the methods described by Southern, *J. Mol. Biol.*, 98:503 (1975). If the gene used as a hybridization probe hybridizes to multiple restriction endonuclease fragments of the genome, that gene is a member of a repertoire of conserved genes.

Immunoglobulins

The immunoglobulins, or antibody molecules, are a large family of molecules that include several types of molecules, such as IgD, IgG, IgA, IgM and IgE. The antibody molecule is typically comprised of two heavy (H) and light (L) chains with both a variable (V) and constant (C) region present on each chain as shown in FIG. 1. Schematic diagrams of human IgG heavy chain and human kappa light chain are shown in FIGS. 2A and 2B, respectively. Several different regions of an immunoglobulin contain conserved sequences useful for isolating an immunoglobulin repertoire. Extensive amino acid and nucleic acid sequence data displaying exemplary conserved sequences is compiled for immunoglobulin molecules by Kabat et al., in *Sequences of Proteins of Immunological Interest*, National Institutes of Health, Bethesda, Md., 1987.

The C region of the H chain defines the particular immunoglobulin type. Therefore the selection of conserved sequences as defined herein from the C region of the H chain results in the preparation of a repertoire of immunoglobulin genes having members of the immunoglobulin type of the selected C region.

The V region of the H or L chain typically comprises four framework (FR) regions each containing relatively lower degrees of variability that includes lengths of conserved sequences. The use of conserved sequences from the FR1 and FR4 (J region) framework regions of the V_H chain is a preferred exemplary embodiment and is described herein in the Examples. Framework regions are typically conserved across several or all immunoglobulin types and thus conserved sequences contained therein are particularly suited for preparing repertoires having several immunoglobulin types.

Major Histocompatibility Complex

The major histocompatibility complex (MHC) is a large genetic locus that encodes an extensive family of proteins that include several classes of molecules referred to as class I, class II or class III MHC molecules. Paul et al., in *Fundamental Immunology*, Raven Press, N.Y., pp. 303–378 (1984).

Class I MHC molecules are a polymorphic group of transplantation antigens representing a conserved family in which the antigen is comprised of a heavy chain and a non-MHC encoded light chain. The heavy chain includes several regions, termed the N, C1, C2, membrane and cytoplasmic regions. Conserved sequences useful in the present invention are found primarily in the N, C1 and C2 regions and are identified as continuous sequences of "invariant residues" in Kabat et al., supra.

Class II MHC molecules comprise a conserved family of polymorphic antigens that participate in immune responsiveness and are comprised of an alpha and a beta chain. The genes coding for the alpha and beta chain each include several regions that contain conserved sequences suitable for producing MHC class II alpha or beta chain repertoires. Exemplary conserved nucleotide sequences include those coding for amino acid residues 26–30 of the A1 region,

residues 161–170 of the A2 region and residues 195–206 of the membrane region, all of the alpha chain. Conserved sequences are also present in the B1, B2 and membrane regions of the beta chain at nucleotide sequences coding for amino acid residues 41–45, 150–162 and 200–209, respectively.

Lymphocyte Receptors and Cell Surface Antigens

Lymphocytes contain several families of proteins on their cell surfaces including the T-cell receptor, Thy-1 antigen and numerous T-cell surface antigens including the antigens defined by the monoclonal antibodies OKT4 (leu3), OKT5/8 (leu2), OKT3, OKT1 (leu1), OKT 11 (leu5) OKT6 and OKT9. Paul, supra at pp. 458–479.

The T-cell receptor is a term used for a family of antigen binding molecules found on the surface of T-cells. The T-cell receptor as a family exhibits polymorphic binding specificity similar to immunoglobulins in its diversity. The mature T-cell receptor is comprised of alpha and beta chains each having a variable (V) and constant (C) region. The similarities that the T-cell receptor has to immunoglobulins in genetic organization and function shows that T-cell receptor contains regions of conserved sequence. Lai et al., *Nature*, 331:543–546 (1988).

Exemplary conserved sequences include those coding for amino acid residues 84–90 of alpha chain, amino acid residues 107–115 of beta chain, and amino acid residues 91–95 and 111–116 of the gamma chain. Kabat et al., supra, p. 279.

Integrins And Adhesions

Adhesive proteins involved in cell attachment are members of a large family of related proteins termed integrins. Integrins are heterodimers comprised of a beta and an alpha subunit. Members of the integrin family include the cell surface glycoproteins platelet receptor GpIIb-IIIa, vitronectin receptor (VnR), fibronectin receptor (FnR) and the leukocyte adhesion receptors LFA-1, Mac-1, Mo-1 and 60.3. Rouslahti et al., *Science*, 238:491–497 (1987). Nucleic acid and protein sequence data demonstrates regions of conserved sequences exist in the members of these families, particularly between the beta chain of GpIIb-IIIa, VnR and FnR, and between the alpha subunit of VnR, Mac-1, LFA-1, FnR and GpIIb-IIIa. Suzuki et al., *Proc. Natl. Acad. Sci. USA*, 83:8614–8618. 1986; Ginsberg et al., *J. Biol. Chem.*, 262:5437–5440, 1987.

Various well known methods can be employed to produce a useful gene repertoire. For instance, V_H and V_L gene repertoires can be produced by isolating V_H- and V_L-coding mRNA from a heterogeneous population of antibody producing cells, i.e., B lymphocytes (B cells), preferably rearranged B cells such as those found in the circulation or spleen of a vertebrate. Rearranged B cells are those in which immunoglobulin gene translocation, i.e., rearrangement, has occurred as evidenced by the presence in the cell of mRNA with the immunoglobulin gene V, D and J region transcripts adjacently located thereon. Typically, the B cells are collected in a 1–100 ml sample of blood which usually contains 10^6 B cells/ml.

In some cases, it is desirable to bias a repertoire for a preselected activity, such as by using as a source of nucleic acid cells (source cells) from vertebrates in any one of various stages of age, health and immune response. For example, repeated immunization of a healthy animal prior to collecting rearranged B cells results in obtaining a repertoire enriched for genetic material producing a receptor of high affinity. Mullinax et al., *Proc. Natl. Acad. Sci. USA*, 87:8095–8099 (1990). Conversely, collecting rearranged B cells from a healthy animal whose immune system has not been recently challenged results in producing a repertoire that is not biased towards the production of high affinity V_H and/or V_L polypeptides.

It should be noted the greater the genetic heterogeneity of the population of cells for which the nucleic acids are obtained, the greater the diversity of the immunological repertoire (comprising V_H- and V_L-coding genes) that will be made available for screening according to the method of the present invention. Thus, cells from different individuals, particularly those having an immunologically significant age difference, and cells from individuals of different strains, races or species can be advantageously combined to increase the heterogeneity (diversity) of a repertoire.

Thus, in one preferred embodiment, the source cells are obtained from a vertebrate, preferably a mammal, which has been immunized or partially immunized with an antigenic ligand (antigen) against which activity is sought, i.e., a preselected antigen. The immunization can be carried out conventionally. Antibody titer in the animal can be monitored to determine the stage of immunization desired, which stage corresponds to the amount of enrichment or biasing of the repertoire desired. Partially immunized animals typically receive only one immunization and cells are collected from those animals shortly after a response is detected. Fully immunized animals display a peak titer, which is achieved with one or more repeated injections of the antigen into the host mammal, normally at 2 to 3 week intervals. Usually three to five days after the last challenge, the spleen is removed and the genetic repertoire of the splenocytes, about 90% of which are rearranged B cells, is isolated using standard procedures. See, *Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, NY. Nucleic acids coding for V_H and V_L polypeptides can be derived from cells producing IgA, IgD, IgE, IgG or IgM, most preferably from IgM and IgG, producing cells.

Methods for preparing fragments of genomic DNA from which immunoglobulin variable region genes can be cloned as a diverse population are well known in the art. See for example Herrmann et al., *Methods In Enzymol.*, 152:180–183, (1987); Frischauf, *Methods In Enzymol.*, 152:183–190 (1987); Frischauf, *Methods In Enzymol.*, 152:190–199 (1987); and DiLella et al., *Methods In Enzymol.*, 152:199–212 (1987). (The teachings of the references cited herein are hereby incorporated by reference.)

The desired gene repertoire can be isolated from either genomic material containing the gene expressing the variable region or the messenger RNA (mRNA) which represents a transcript of the variable region. The difficulty in using the genomic DNA from other than non-rearranged B lymphocytes is in juxtaposing the sequences coding for the variable region, where the sequences are separated by introns. The DNA fragment(s) containing the proper exons must be isolated, the introns excised, and the exons then spliced in the proper order and in the proper orientation. For the most part, this will be difficult, so that the alternative technique employing rearranged B cells will be the method of choice because the V, D and J immunoglobulin gene regions have translocated to become adjacent, so that the sequence is continuous (free of introns) for the entire variable regions.

Where mRNA is utilized the cells will be lysed under RNase inhibiting conditions. In one embodiment, the first step is to isolate the total cellular mRNA. Poly A+ mRNA can then be selected by hybridization to an oligo-dT cellulose column. The presence of mRNAs coding for the heavy and/or light chain polypeptides can then be assayed by hybridization with DNA single strands of the appropriate

genes. Conveniently, the sequences coding for the constant portion of the V_H and V_L can be used as polynucleotide probes, which sequences can be obtained from available sources. See for example, Early and Hood, *Genetic Engineering*, Setlow and Hollaender, eds., Vol. 3, Plenum Publishing Corporation, NY, (1981), pages 157–188; and Kabat et al., *Sequences of Immunological Interest*, National Institutes of Health, Bethesda, Md., (1987).

In preferred embodiments, the preparation containing the total cellular mRNA is first enriched for the presence of V_H and/or V_L coding mRNA. Enrichment is typically accomplished by subjecting the total mRNA preparation or partially purified mRNA product thereof to a primer extension reaction employing a polynucleotide synthesis primer as described herein. Exemplary methods for producing V_H and V_L gene repertoires using polynucleotide synthesis primers are described in PCT Application No. PCT/US 90/02836 (International Publication No. WO 90/14430). Particularly preferred methods for producing a gene repertoire rely on the use of preselected oligonucleotides as primers in a polymerase chain reaction (PCR) to form PCR reaction products as described herein.

In preferred embodiments, isolated B cells are immunized in vitro against a preselected antigen. In vitro immunization is defined as the clonal expansion of epitope-specific B cells in culture, in response to antigen stimulation. The end result is to increase the frequency of antigen-specific B cells in the immunoglobulin repertoire, and thereby decrease the number of clones in an expression library that must be screened to identify a clone expressing an antibody of the desired specificity. The advantage of in vitro immunization is that human monoclonal antibodies can be generated against a limitless number of therapeutically valuable antigens, including toxic or weak immunogens. For example, antibodies specific for the polymorphic determinants of tumor-associated antigens, rheumatoid factors, and histocompatibility antigens can be produced, which can not be elicited in immunized animals. In addition, it may be possible to generate immune responses which are normally suppressed in vivo.

In vitro immunization can be used to give rise to either a primary or secondary immune response. A primary immune response, resulting from first time exposure of a B cell to an antigen, results in clonal expansion of epitope-specific cells and the secretion of IgM antibodies with low to moderate apparent affinity constants (10^6–10^8 M^{-1}). Primary immunization of human splenic and tonsillar lymphocytes in culture can be used to produce monoclonal antibodies against a variety of antigens, including cells, peptides, macromolecule, haptens, and tumor-associated antigens. Memory B cells from immunized donors can also be stimulated in culture to give rise to a secondary immune response characterized by clonal expansion and the production of high affinity antibodies (>10^9 M^{-1}) of the IgG isotype, particularly against viral antigens by clonally expanding sensitized lymphocytes derived from seropositive individuals.

In one embodiment, peripheral blood lymphocytes are depleted of various cytolytic cells that appear to down-modulate antigen-specific B cell activation. When lysosome-rich subpopulations (natural killer cells, cytotoxic and suppressor T cells, monocytes) are first removed by treatment with the lysosmotropic methyl ester of leucine, the remaining cells (including B cells, T helper cells, accessory cells) respond antigen-specifically during in vitro immunization. The lymphokine requirements for inducing antibody production in culture are satisfied by a culture supernatant from activated, irradiated T cells.

In addition to in vitro immunization, cell panning (immunoaffinity absorption) can be used to further increase the frequency of antigen-specific B cells. Techniques for selecting B cell subpopulations via solid-phase antigen binding are well established. Panning conditions can be optimized to selectively enrich for B cells which bind with high affinity to a variety of antigens, including cell surface proteins. Panning can be used alone, or in combination with in vitro immunization to increase the frequency of antigen-specific cells above the levels which can be obtained with either technique alone. Immunoglobulin expression libraries constructed from enriched populations of B cells are biased in favor of antigen-specific antibody clones, and thus, enabling identification of clones with the desired specificities from smaller, less complex libraries.

3. Preparation of Polynucleotide Primers

The term "polynucleotide" as used herein in reference to primers, probes and nucleic acid fragments or segments to be synthesized by primer extension is defined as a molecule comprised of two or more deoxyribonucleotide or ribonucleotides, preferably more than 3. Its exact size will depend on many factors, which in turn depends on the ultimate conditions of use.

The term "primer" as used herein refers to a polynucleotide whether purified from a nucleic acid restriction digest or produced synthetically, which is capable of acting as a point of initiation of nucleic acid synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase, reverse transcriptase and the like, and at a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency, but may alternatively be in double stranded form. If double stranded, the primer is first treated to separate it from its complementary strand before being used to prepare extension products. Preferably, the primer is a polydeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agents for polymerization. The exact lengths of the primers will depend on may factors, including temperature and the source of primer. For example, depending on the complexity of the target sequence, a polynucleotide primer typically contains 15 to 25 or more nucleotides, although it can contain fewer nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with template.

The primers used herein are selected to be "substantially" complementary to the different strands of each specific sequence to be synthesized or amplified. This means that the primer must be sufficiently complementary to non-randomly hybridize with its respective template strand. Therefore, the primer sequence may or may not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment can be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Such non-complementary fragments typically code for an endonuclease restriction site. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided the primer sequence has sufficient complementarily with the sequence of the strand to be synthesized or amplified to non-randomly hybridize therewith and thereby form an extension product under polynucleotide synthesizing conditions.

Primers of the present invention may also contain a DNA-dependent RNA polymerase promoter sequence or its

complement. See for example. Krieg et al., *Nucl. Acids Res.*, 12:7057–70 (1984); Studier et al., *J. Mol. Biol.*, 189:113–130 (1986); and *Molecular Cloning: A Laboratory Manual, Second Edition*, Maniatis et al., eds., Cold Spring Harbor, N.Y. (1989).

When a primer containing a DNA-dependent RNA polymerase promoter is used the primer is hybridized to the polynucleotide strand to be amplified and the second polynucleotide strand of the DNA-dependent RNA polymerase promoter is completed using an inducing agent such as *E. coli* DNA polymerase I, or the Klenow fragment of *E. coli* DNA polymerase. The starting polynucleotide is amplified by alternating between the production of an RNA polynucleotide and DNA polynucleotide.

Primers may also contain a template sequence or replication initiation site for a RNA-directed RNA polymerase. Typical RNA-directed RNA polymerase include the QB replicase described by Lizardi et al., *Biotechnology*, 6:1197–1202 (1988) RNA-directed polymerases produce large numbers of RNA strands from a small number of template RNA strands that contain a template sequence or replication initiation site. These polymerases typically give a one million-fold amplification of the template strand as has been described by Kramer et al., *J. Mol. Biol.*, 89:719–736 (1974).

The polynucleotide primers can be prepared using any suitable method, such as, for example, the phosphotriester or phosphodiester methods see Narang et al., *Meth. Enzymol.*, 68:90, (1979); U.S. Pat. No. 4,356,270; and Brown et al., *Meth. Enzymol.*, 68:109, (1979).

The choice of a primer's nucleotide sequence depends on factors such as the distance on the nucleic acid from the region coding for the desired receptor, its hybridization site on the nucleic acid relative to any second primer to be used, the number of genes in the repertoire it is to hybridize to, and the like.

a. Primers for Producing Immunoglobulin Gene Repertoires

V_H and V_L gene repertoires can be separately prepared prior to their utilization in the present invention. Repertoire preparation is typically accomplished by primer extension, preferably by primer extension in a polymerase chain reaction (PCR) format.

To produce a repertoire of V_H-coding DNA homologs by primer extension, the nucleotide sequence of a primer is selected to hybridize with a plurality of immunoglobulin heavy chain genes at a site substantially adjacent to the V_H-coding region so that a nucleotide sequence coding for a functional (capable of binding) polypeptide is obtained. To hybridize to a plurality of different V_H-coding nucleic acid strands, the primer must be a substantial complement of a nucleotide sequence conserved among the different strands. Such sites include nucleotide sequences in the constant region, any of the variable region framework regions, preferably the third framework region, leader region, promoter region, J region and the like.

If the repertoires of V_H-coding and V_L-coding DNA homologs are to be produced by (PCR) amplification, two primers, i.e., a PCR primer pair, must be used for each coding strand of nucleic acid to be amplified. The first primer becomes part of the nonsense (minus or complementary) strand and hybridizes to a nucleotide sequence conserved among V_H (plus or coding) strands within the repertoire. To produce V_H coding DNA homologs, first primers are therefore chosen to hybridize to (i.e. be complementary to) conserved regions within the J region, CH1 region, hinge region, CH2 region, or CH3 region of immunoglobulin genes and the like. To produce a V_L coding DNA homolog, first primers are chosen to hybridize with (i.e. be complementary to) a conserved region within the J region or constant region of immunoglobulin light chain genes and the like. Second primers become part of the coding (plus) strand and hybridize to a nucleotide sequence conserved among minus strands. To produce the V_H-coding DNA homologs, second primers are therefore chosen to hybridize with a conserved nucleotide sequence at the 5' end of the V_H-coding immunoglobulin gene such as in that area coding for the leader or first framework region. It should be noted that in the amplification of both V_H- and V_L-coding DNA homologs the conserved 5' nucleotide sequence of the second primer can be complementary to a sequence exogenously added using terminal deoxynucleotidyl transferase as described by Loh et al., *Science*, 243:217–220 (1989). One or both of the first and second primers can contain a nucleotide sequence defining an endonuclease recognition site. The site can be heterologous to the immunoglobulin gene being amplified and typically appears at or near the 5' end of the primer.

When present, the restriction site-defining portion is typically located in a 5'-terminal non-priming portion of the primer. The restriction site defined by the first primer is typically chosen to be one recognized by a restriction enzyme that does not recognize the restriction site defined by the second primer, the objective being to be able to produce a DNA molecule having cohesive termini that are non-complementary to each other and thus allow directional insertion into a vector.

In one embodiment, the present invention utilizes a set of polynucleotides that form primers having a priming region located at the 3'-terminus of the primer. The priming region is typically the 3'-most (3'-terminal) 15 to 30 nucleotide bases. The 3'-terminal priming portion of each primer is capable of acting as a primer to catalyze nucleic acid synthesis, i.e., initiate a primer extension reaction off its 3' terminus. One or both of the primers can additionally contain a 5'-terminal (5'-most) non-priming portion, i.e., a region that does not participate in hybridization to repertoire template.

In PCR, each primer works in combination with a second primer to amplify a target nucleic acid sequence. The choice of PCR primer pairs for use in PCR is governed by considerations as discussed herein for producing gene repertoires. That is, the primers have a nucleotide sequence that is complementary to a sequence conserved in the repertoire. Useful V_H and V_L priming sequences are shown in Tables 5 and 6, herein below.

4. Polymerase Chain Reaction to Produce Gene Repertoires

The strategy used for cloning the V_H and V_L genes contained within a repertoire will depend, as is well known in the art, on the type, complexity, and purity of the nucleic acids making up the repertoire. Other factors include whether or not the genes are contained in one or a plurality of repertoires and whether or not they are to be amplified and/or mutagenized.

The V_H- and V_L-coding gene repertoires are comprised of polynucleotide coding strands, such as mRNA and/or the sense strand of genomic DNA. If the repertoire is in the form of double stranded genomic DNA, it is usually first denatured, typically by melting, into single strands. A repertoire is subjected to a PCR reaction by treating (contacting) the repertoire with a PCR primer pair, each member of the pair having a preselected nucleotide sequence. The PCR primer pair is capable of initiating primer extension reactions by hybridizing to nucleotide sequences, preferably at least about 10 nucleotides in length

and more preferably at least about 20 nucleotides in length, conserved within the repertoire. The first primer of a PCR primer pair is sometimes referred to herein as the "sense primer" because it hybridizes to the coding or sense strand of a nucleic acid. In addition, the second primer of a PCR primer pair is sometimes referred to herein as the "anti-sense primer" because it hybridizes to a non-coding or anti-sense strand of a nucleic acid, i.e., a strand complementary to a coding strand.

The PCR reaction is performed by mixing the PCR primer pair, preferably a predetermined amount thereof, with the nucleic acids of the repertoire, preferably a predetermined amount thereof, in a PCR buffer to form a PCR reaction admixture. The admixture is maintained under polynucleotide synthesizing conditions for a time period, which is typically predetermined, sufficient for the formation of a PCR reaction product, thereby producing a plurality of different V_H-coding and/or V_L-coding DNA homologs.

A plurality of first primer and/or a plurality of second primers can be used in each amplification, e.g., one species of first primer can be paired with a number of different second primers to form several different primer pairs. Alternatively, an individual pair of first and second primers can be used. In any case, the amplification products of amplifications using the same or different combinations of first and second primers can be combined to increase the diversity of the gene library.

In another strategy, the object is to clone the V_H- and/or V_L-coding genes from a repertoire by providing a polynucleotide complement of the repertoire, such as the anti-sense strand of genomic dsDNA or the polynucleotide produced by subjecting mRNA to a reverse transcriptase reaction. Methods for producing such complements are well known in the art.

The PCR reaction is performed using any suitable method. Generally it occurs in a buffered aqueous solution, i.e., a PCR buffer, preferably at a pH of 7–9, most preferably about 8. Preferably, a molar excess (for genomic nucleic acid, usually about 10^6:1 primer:template) of the primer is admixed to the buffer containing the template strand. A large molar excess is preferred to improve the efficiency of the process.

The PCR buffer also contains the deoxyribonucleotide triphosphates dATP, dCTP, dGTP, and dTTP and a polymerase, typically thermostable, all in adequate amounts for primer extension (polynucleotide synthesis) reaction. The resulting solution (PCR admixture) is heated to about 90° C. –100° C. for about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period the solution is allowed to cool to 54° C., which is preferable for primer hybridization. The synthesis reaction may occur at from room temperature up to a temperature above which the polymerase (inducing agent) no longer functions efficiently. Thus, for example, if DNA polymerase is used as inducing agent, the temperature is generally no greater than about 40° C. An exemplary PCR buffer comprises the following: 50 mM KCl; 10 mM Tris-HCl; pH 8.3; 1.5 mM $MgCl_2$; 0.001% (wt/vol) gelatin, 200 μM dATP; 200 μM dTTP; 200 μM dCTP; 200 μM dGTP; and 2.5 units *Thermus aquaticus* DNA polymerase I (U.S. Pat. No. 4,889,818) per 100 microliters of buffer.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase, and other enzymes, including heat-stable enzymes, which will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths. There may be inducing agents, however, which initiate synthesis at the 5' end and proceed in the above direction, using the same process as described above.

The inducing agent also may be a compound or system which will function to accomplish the synthesis of RNA primer extension products, including enzymes. In preferred embodiments, the inducing agent may be a DNA-dependent RNA polymerase such as T7 RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase. These polymerases produce a complementary RNA polynucleotide. The high turn over rate of the RNA polymerase amplifies the starting polynucleotide as has been described by Chamberlin et al., *The Enzymes*, ed. P. Boyer, PP. 87–108, Academic Press, New York (1982). Another advantage of T7 RNA polymerase is that mutations can be introduced into the polynucleotide synthesis by replacing a portion of cDNA with one or more mutagenic oligodeoxynucleotides (polynucleotides) and transcribing the partially-mismatched template directly as has been previously described by Joyce et al., *Nuc. Acid Res.*, 17:711–722 (1989). Amplification systems based on transcription have been described by Gingeras et al., in *PCR Protocols, A Guide to Methods and Applications*, pp 245–252, Academic Press, Inc., San Diego, Calif. (1990).

If the inducing agent is a DNA-dependent RNA polymerase and therefore incorporates ribonucleotide triphosphates, sufficient amounts of ATP, CTP, GTP and UTP are admixed to the primer extension reaction admixture and the resulting solution is treated as described above.

The newly synthesized strand and its complementary nucleic acid strand form a double-stranded molecule which can be used in the succeeding steps of the process.

The first and/or second PCR reactions discussed above can advantageously be used to incorporate into the receptor a preselected epitope useful in immunologically detecting and/or isolating a receptor. This is accomplished by utilizing a first and/or second polynucleotide synthesis primer or expression vector to incorporate a predetermined amino acid residue sequence into the amino acid residue sequence of the receptor.

After producing V_H- and V_L-coding DNA homologs for a plurality of different V_H- and V_L-coding genes within the repertoires, the DNA molecules are typically further amplified. While the DNA molecules can be amplified by classic techniques such as incorporation into an autonomously replicating vector, it is preferred to first amplify the molecules by subjecting them to a polymerase chain reaction (PCR) prior to inserting them into a vector. PCR is typically carried out by thermocycling i.e., repeatedly increasing and decreasing the temperature of a PCR reaction admixture within a temperature range whose lower limit is about 10° C. to about 40° C. and whose upper limit is about 90° C. to about 100° C. The increasing and decreasing can be continuous, but is preferably phasic with time periods of relative temperature stability at each of temperatures favoring polynucleotide synthesis, denaturation and hybridization.

PCR amplification methods are described in detail in U.S. Pat. Nos. 4,683,192, 4,683,202, 4,800,159, and 4,965,188.

and at least in several texts including "PCR Technology: Principles and Applications for DNA Amplification", H. Erlich, ed., Stockton Press, New York (1989); and "PCR Protocols: A Guide to Methods and Applications", Innis et al., eds., Academic Press, San Diego, Calif. (1990).

In preferred embodiments only one pair of first and second primers is used per amplification reaction. The amplification reaction products obtained from a plurality of different amplifications, each using a plurality of different primer pairs, are then combined.

However, the present invention also contemplates DNA homolog production via co-amplification (using two pairs of primers), and multiplex amplification (using up to about 8, 9 or 10 primer pairs).

In preferred embodiments, the PCR process is used not only to produce a library of DNA molecules, but also to induce mutations within the library or to create diversity from a single parental clone and thereby provide a library having a greater heterogeneity. First, it should be noted that the PCR process itself is inherently mutagenic due to a variety of factors well known in the art. Second, in addition to the mutation inducing variations described in the above referenced U.S. Pat. No. 4,683,195, other mutation inducing PCR variations can be employed. For example, the PCR reaction admixture, can be formed with different amounts of one or more of the nucleotides to be incorporated into the extension product. Under such conditions, the PCR reaction proceeds to produce nucleotide substitutions within the extension product as a result of the scarcity of a particular base. Similarly, approximately equal molar amounts of the nucleotides can be incorporated into the initial PCR reaction admixture in an amount to efficiently perform X number of cycles, and then cycling the admixture through a number of cycles in excess of X, such as, for instance, 2X. Alternatively, mutations can be induced during the PCR reaction by incorporating into the reaction admixture nucleotide derivatives such as inosine, not normally found in the nucleic acids of the repertoire being amplified. During subsequent in vivo amplification, the nucleotide derivative will be replaced with a substitute nucleotide thereby inducing a point mutation.

5. Linear DNA Expression Vectors

A DNA expression vector for use in a method of the invention for producing a library of DNA molecules is a linearized DNA molecule as described before having two (upstream and downstream) cohesive termini adapted for directional ligation to a polypeptide gene.

A linear DNA expression vector is typically prepared by restriction endonuclease digestion of a circular DNA expression vector of this invention to cut at two preselected restriction sites within the sequence of nucleotides of the vector adapted for directional ligation to produce a linear DNA molecule having the required cohesive termini that are adapted for direction ligation. Directional ligation refers to the presence of two (a first and second) cohesive termini on a vector, or on the insert DNA molecule to be ligated into the vector selected, so that the termini on a single molecule are not complementary. A first terminus of the vector is complementary to a first terminus of the insert, and the second terminus of the vector is complementary to the second terminus of the insert.

6. Ligation Reactions to Produce Gene Libraries

In preparing a library of DNA molecules of this invention, a ligation admixture is prepared as described above, and the admixture is subjected to ligation conditions for a time period sufficient for the admixed repertoire of polypeptide genes to ligate (become operatively linked) to the plurality of DNA expression vectors to form the library.

Ligation conditions are conditions selected to favor a ligation reaction wherein a phosphodiester bond is formed between adjacent 3' hydroxyl and 5' phosphoryl termini of DNA. The ligation reaction is preferably catalyzed by the enzyme T4 DNA ligase. Ligation conditions can vary in time, temperature, concentration of buffers, quantities of DNA molecules to be ligated, and amounts of ligase, as is well known. Preferred ligation conditions involve maintaining the ligation admixture at 4 degrees Centigrade (4° C.) to 12° C. for 1 to 24 hours in the presence of 1 to 10 units of T4 DNA ligase per milliliter (ml) and about 1 to 2 micrograms (ug) of DNA. Ligation buffer in a ligation admixture typically contains 0.5M Tris-HCl (pH 7.4), 0.01M $MgCl_2$, 0.01M dithiothrietol, 1 mM spermidine, 1 mM ATP and 0.1 mg/ml bovine serum albumin (BSA). Other ligation buffers can also be used.

Exemplary ligation reactions are described in Example 2.

7. Preparation of Dicistronic Gene Libraries

In a particularly preferred embodiment, the present invention contemplates methods for the preparation of a library of dicistronic DNA molecules. A dicistronic DNA molecule is a single DNA molecule having the capacity to express two separate polypeptides from two separate cistrons. In preferred embodiments, the two cistrons are operatively linked at relative locations on the DNA molecule such that both cistrons are under the transcriptional control of a single promoter. Each dicistronic molecule is capable of expressing first and second polypeptides from first and second cistrons, respectively, that can form, in a suitable host, a heterodimeric receptor on the surface of a filamentous phage particle.

The method for producing a library of dicistronic DNA molecules comprises the steps of:

(a) Forming a first ligation admixture by combining in a ligation buffer:

(i) a repertoire of first polypeptide genes in the form of dsDNA, each having cohesive termini adapted for directional ligation, and

(ii) a plurality of DNA expression vectors in linear form, each having upstream and downstream first cohesive termini that are (a) adapted for directionally receiving the first polypeptide genes in a common reading frame, and (b) operatively linked to respective upstream and downstream translatable DNA sequences. The upstream translatable DNA sequence encodes a pelB secretion signal, the downstream translatable DNA sequence encodes a filamentous phage coat protein membrane anchor, and translatable DNA sequences are operatively linked to respective upstream and downstream DNA expression control sequences.

(b) Subjecting the admixture to ligation conditions for a time period sufficient to operatively link the first polypeptide genes to the vectors and produce a plurality of circular DNA molecules each having a first cistron for expressing the first polypeptide.

(c) Treating the plurality of circular DNA molecules under DNA cleavage conditions to produce a plurality of DNA expression vectors in linear form that each have upstream and downstream second cohesive termini that are (i) adapted for directionally receiving a repertoire of second polypeptide genes in a common reading frame, and (ii) operatively linked to respective upstream and downstream DNA sequences. The upstream DNA sequence is a translatable sequence encoding a secretion signal, the downstream DNA sequence has at least one stop codon in the reading frame, and the translat-

able DNA sequence is operatively linked to a DNA expression control sequence.

(d) Forming a second ligation admixture by combining in a ligation buffer:

(i) the plurality of DNA expression vectors formed in step (c). and

(ii) the repertoire of second polypeptide genes in the form of dsDNA. each having cohesive termini adapted for directional ligation to the plurality of DNA expression vectors; and

(e) Subjecting the second admixture to ligation conditions for a time period sufficient to operatively link the second polypeptide genes to said vectors and produce a plurality of circular DNA molecules each having the second cistron for expressing the second polypeptide. thereby forming the library. In preferred embodiments a secretion signal is a pelB secretion signal. and the membrane anchor is derived from cpVIII as described herein.

DNA expression vectors useful for practicing the above method are the dicistronic expression vectors described in greater detail before.

In practicing the method of producing a library of dicistronic DNA molecules. it is preferred that the upstream and downstream first cohesive termini do not have the same nucleotide sequences as the upstream and downstream second cohesive termini. In this embodiment. the treating step (c) to linearize the circular DNA molecules typically involves the use of restriction endonucleases that are specific for producing said second termini. but do not cleave the circular DNA molecule at the sites that formed the first termini. Exemplary and preferred first and second termini are the termini defined by cleavage of pCBAK8 with Xho I and Spe I to form the upstream and downstream first termini. and defined by cleavage of pCBAK8 with Sac I and Xba I to form the upstream and downstream second termini. In this embodiment. other pairs of cohesive termini can be utilized at the respective pairs of first and second termini. so long as the four termini are each distinct. non-complementary termini.

Methods of treating the plurality of circular DNA molecules under DNA cleavage conditions to form linear DNA molecules are generally well known and depend on the nucleotide sequence to be cleaved and the mechanism for cleavage. Preferred treatments involve admixing the DNA molecules with a restriction endonuclease specific for a endonuclease recognition site at the desired cleavage location in an amount sufficient for the restriction endonuclease to cleave the DNA molecule. Buffers. cleavage conditions. and substrate concentrations for restriction endonuclease cleavage are well known and depend on the particular enzyme utilized. Exemplary restriction enzyme cleavage conditions are described in Example 2.

F. Phage Libraries

The present invention contemplates a library of DNA molecules that each encode a fusion protein of this invention where the library is in the form of a population of different filamentous phage particles each containing one of the different rDNA molecules. By different ,DNA molecule is meant rDNA molecules differing in nucleotide base sequence encoding a polypeptide of this invention.

Thus. a phage library is a population of filamentous phage. preferably f1. fd or M13 filamentous phage. each phage having packaged inside the particle a rDNA expression vector of this invention. A preferred library is comprised of phage particles containing DNA molecules that encode at least 10^6. preferably 10^7 and more preferably 10^{8-9} different fusion proteins of this invention. By different fusion proteins is meant fusion proteins differing in amino acid residue sequence. Where the packaged expression vector encodes first and second polypeptides of an autogenously assembling receptor. e.g. V_H and V_L polypeptides that form a Fab. the library can also be characterized as containing or expressing a multiplicity of receptor specificities. Thus. preferred libraries express at least 10^5. preferably at least 10^6 and more preferably at least 10^7 different receptors. such as different antibodies. T cell receptors. integrins and the like.

As described herein. a particular advantage of a filamentous phage in the present invention is that the DNA molecule present in the phage particle and encoding one or both of the members of the heterodimeric receptor can be segregated from other DNA molecules present in the library on the basis of the presence of the particular expressed fusion protein the surface of the phage particle.

Isolation (segregation) of a DNA molecule encoding one or both members of a heterodimeric receptor is conducted by segregation of the filamentous phage particle containing the gene or genes of interest away from the population of other phage particles comprising the library. Segregation of phage particles involves the physical separation and propagation of individual phage particles away from other particles in the library. Methods for physical separation of filamentous phage particles to produce individual particles. and the propagation of the individual particles to form populations of progeny phage derived from the individual segregated particle are well known in the filamentous phage arts.

A preferred separation method involves the identification of the expressed heterodimer on the surface of the phage particle by means of a ligand binding specificity between the phage particle and a preselected ligand. Exemplary and preferred is the use of "panning" methods whereby a suspension of phage particles is contacted with a solid phase ligand (antigen) and allowed to specifically bind (or immunoreact where the heterodimer includes an immunoglobulin variable domain). After binding. non-bound particles are washed off the solid phase. and the bound phage particles are those that contain ligand-specific heterodimeric receptor (heterodimer) on their surface. The bound particles can then be recovered by elution of the bound particle from the solid phase. typically by the use of aqueous solvents having high ionic strength sufficient to disrupt the receptor-ligand binding interaction.

An alternate method for separating a phage particle based on the ligand specificity of the surface-expressed heterodimer from a population of particles is to precipitate the phage particles from the solution phase by crosslinkage with the ligand. An exemplary and preferred crosslinking and precipitation method is described in detail in Example 4c.

The use of the above particle segregation methods provides a means for screening a population of filamentous phage particles present in a phage library of this invention. As applied to a phage library. screening can be utilized to enrich the library for one or more particles that express a heterodimer having a preselected ligand binding specificity. Where the library is designed to contain multiple species of heterodimers that all have some detectable measure of ligand binding activity. but differ in protein structure. antigenicity. ligand binding affinity or avidity. and the like. the screening methods can be utilized sequentially to first produce a library enriched for a preselected binding specificity. and then to produce a second library further

enriched by further screening comprising one or more isolated phage particles. Methods for measuring ligand binding activities, antigenicity and the like interactions between a ligand and a receptor are generally well known and are not discussed further as they are not essential features of the present invention.

Thus, in one embodiment, a phage library is a population of particles enriched for a preselected ligand binding specificity.

In another embodiment, a phage library comprises a population of particles wherein each particle contains at least one fusion protein of this invention on the surface of the phage particle. The actual amount of fusion protein present on the surface of a phage particle depends, in part, on the choice of coat protein membrane anchor present in the fusion protein. Where the anchor is derived from cpIII, there are typically about 1 to 4 fusion proteins per phage particle. Where the anchor is derived from the more preferred cpVIII, there is the potential for hundreds of fusion proteins on the particle surface depending on the growth conditions and other factors as discussed herein. Preferably, a phage particle in a library contains from about 10 to about 500 cpVIII-derived fusion proteins on the surface of each particle, and more preferably about 20 to 50 fusion proteins per particle. Exemplary amounts of surface fusion protein are shown by the electron micrographs described in Example 4a that describe particles having about 20 to 24 cpVIII-derived fusion proteins per particle.

In another embodiment, the present invention contemplates a population of phage particles that are the progeny of a single particle, and therefor all express the same heterodimer on the particle surface. Such a population of phage are homogeneous and clonally derived, and therefore provide a source for expressing large quantities of a particular fusion protein. An exemplary clonally homogeneous phage population is described in Example 4.

A filamentous phage particle in a library of this invention is produced by standard filamentous phage particle preparation methods and depends on the presence in a DNA expression vector of this invention of a filamentous phage origin of replication as described herein to provide the signals necessary for (1) production of a single-stranded filamentous phage replicative form and (2) packaging of the replicative form into a filamentous phage particle. Such a DNA molecule can be packaged when present in a bacterial cell host upon introduction of genetic complementation to provide the filamentous phage proteins required for production of infectious phage particles. A typical and preferred method for genetic complementation is to infect a bacterial host cell containing a DNA expression vector of this invention with a helper filamentous phage, thereby providing the genetic elements required for phage particle assembly. Exemplary helper rescue methods are described herein at Example 2, and described by Short et al.: *Nuc. Acids Res.*, 16:7583–7600 (1988).

The level of heterodimeric receptor captured on the surface of a filamentous phage particle during the process of phage particle extrusion from the host cell can be controlled by a variety of means. In one embodiment, the levels of fusion proteins are controlled by the use of strong promoters in the first and second cistrons for expressing the polypeptides, such that transcription of the fusion protein cistrons occurs at a relative rate greater than the rate of transcription of the cpVIII gene on the helper phage. In another embodiment, the helper phage can have an amber mutation in the gene for expressing cpVIII, such that less wild-type cpVIII is transcribed in the host cell than fusion proteins, thereby leading to increased ratios of fusion protein compared to cpVIII during the extrusion process.

In another embodiment, the amount of heterodimeric receptor on the phage particle surface can be controlled by controlling the timing between expression of fusion proteins and the superinfection by helper phage. After introduction of the expression vector, longer delay times before the addition of helper phage will allow for increased accumulation of the fusion proteins in the host cell.

EXAMPLES

The following examples are intended to illustrate, but not limit, the scope of the invention.

1. Construction of a Dicistronic Expression Vector for Producing a Heterodimeric Receptor on Phage Particles

To obtain a vector system for generating a large number of Fab antibody fragments that can be screened directly, expression libraries in bacteriophage Lambda have previously been constructed as described in Huse et al., *Science*, 246:1275–1281 (1989). These systems did not contain design features that provide for the expressed Fab to be targeted to the surface of a filamentous phage particle.

The main criterion used in choosing a vector system was the necessity of generating the largest number of Fab fragments which could be screened directly. Bacteriophage Lambda was selected as the starting point to develop an expression vector for three reasons. First, in vitro packaging of phage DNA was the most efficient method of reintroducing DNA into host cells. Second, it was possible to detect protein expression at the level of single phage plaques. Finally, the screening of phage libraries typically involved less difficulty with nonspecific binding. The alternative, plasmid cloning vectors, are only advantageous in the analysis of clones after they have been identified. This advantage was not lost in the present system because of the use of a dicistronic expression vector such as pCombVIII, thereby permitting a plasmid containing the heavy chain, light chain, or Fab expressing inserts to be excised.

a. Construction of Dicistronic Expression Vector pCOMB

(i) Preparation of Lambda Zap™ II

Lambda Zap™ II is a derivative of the original Lambda Zap (ATCC #40.298) that maintains all of the characteristics of the original Lambda Zap including 6 unique cloning sites, fusion protein expression, and the ability to rapidly excise the insert in the form of a phagemid (Bluescript SK–), but lacks the SAM 100 mutation, allowing growth on many Non-Sup F strains, including XL1-Blue. The Lambda Zap™ II was constructed as described in Short et al., *Nuc. Acids Res.*, 16:7583–7600, 1988, by replacing the Lambda S gene contained in a 4254 base pair (bp) DNA fragment produced by digesting Lambda Zap with the restriction enzyme Nco I. This 4254 bp DNA fragment was replaced with the 4254 bp DNA fragment containing the Lambda S gene isolated from Lambda gt10 (ATCC #40.179) after digesting the vector with the restriction enzyme Nco I. The 4254 bp DNA fragment isolated from lambda gt10 was ligated into the original Lambda Zap vector using T4 DNA ligase and standard protocols such as those described in *Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley and Sons, NY, 1987, to form Lambda Zap™ II.

(ii) Preparation of Lambda Hc2

To express a plurality of V_H-coding DNA homologs in an *E. coli* host cell, a vector designated Lambda Hc2 was constructed. The vector provided the following: the capacity to place the V_H-coding DNA homologs in the proper reading frame; a ribosome binding site as described by Shine et al., *Nature*, 254:34, 1975; a leader sequence directing the expressed protein to the periplasmic space designated the pelB secretion signal; a polynucleotide sequence that coded for a known epitope (epitope tag); and also a polynucleotide that coded for a spacer protein between the V_H-coding DNA homolog and the polynucleotide coding for the epitope tag. Lambda Hc2 has been previously described by Huse et al., *Science*, 246:1275–1281 (1989).

To prepare Lambda Hc2, a synthetic DNA sequence containing all of the above features was constructed by designing single stranded polynucleotide segments of 20–40 bases that would hybridize to each other and form the double stranded synthetic DNA sequence shown in FIG. 3. The individual single-stranded polynucleotide segments are shown in Table 3.

TABLE 3

SEQ. ID. NO.		
(22)	N1)	5' GGCCGCAAATTCTATTTCAAGGAGACAGTCAT 3'
(23)	N2)	5' AATGAAATACCTATTGCCTACGGCAGCCGCTGGATT 3'
(24)	N3)	5' GTTATTACTCGCTGCCCAACCAGCCATGGCCC 3'
(25)	N6)	5' CAGTTTCACCTGGGCCATGGCTGGTTGGG 3'
(26)	N7)	5' CAGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAG 3'
(27)	N8)	5' GTATTTCATTATGACTGTCTCCTTGAAATAGAATTTGC 3'
(28)	N9-4)	5' AGGTGAAACTGCTCGAGATTTCTAGACTAGTTACCCGTAC 3'
(29)	N10-5)	5' CGGAACGTCGTACGGGTAACTAGTCTAGAAATCTCGAG 3'
(30)	N11)	5' GACGTTCCGGACTACGGTTCTTAATAGAATTCG 3'
(31)	N12)	5' TCGACGAATTCTATTAAGAACCGTAGTC 3'

Polynucleotides N2, N3, N9-4, N11, N10-5, N6, N7 and N8 (Table 3) were kinased by adding 1 μl of each polynucleotide 0.1 micrograms/microliter (μg/μl) and 20 units of T_4 polynucleotide kinase to a solution containing 70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 10 mM beta-mercaptoethanol, 500 micrograms per milliliter (μg/ml) bovine serum albumin (BSA). The solution was maintained at 37 degrees Centigrade (37° C.) for 30 minutes and the reaction stopped by maintaining the solution at 65° C. for 10 minutes. The two end polynucleotides, 20 ng of polynucleotides N1 and polynucleotides N12, were added to the above kinasing reaction solution together with 1/10 volume of a solution containing 20.0 mM Tris-HCl, pH 7.4, 2.0 mM $MgCl_2$ and 50.0 mM NaCl. This solution was heated to 70° C. for 5 minutes and allowed to cool to room temperature, approximately 25° C., over 1.5 hours in a 500 ml beaker of water. During this time period all 10 polynucleotides annealed to form the double stranded synthetic DNA insert shown in FIG. 3. The individual polynucleotides were covalently linked to each other to stabilize the synthetic DNA insert by adding 40 μl of the above reaction to a solution containing 50 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 1 mM DTT, 1 mM adenosine triphosphate (ATP) and 10 units of T4 DNA ligase. This solution was maintained at 37° C. for 30 minutes and then the T4 DNA ligase was inactivated by maintaining the solution at 65° C. for 10 minutes. The end polynucleotides were kinased by mixing 52 μl of the above reaction, 4 μl of a solution containing 10 mM ATP and 5 units of T4 polynucleotide kinase. This solution was maintained at 37° C. for 30 minutes and then the T4 polynucleotide kinase was inactivated by maintaining the solution at 65° C. for 10 minutes.

The completed synthetic DNA insert was ligated directly into the Lambda Zap™ II vector described in Example 1a(i) that had been previously digested with the restriction enzymes, Not I and Xho I. The ligation mixture was packaged according to the manufacture's instructions using Gigapack II Gold packing extract available from Stratagene, La Jolla, Calif. The packaged ligation mixture was plated on XL1-Blue cells (Stratagene). Individual lambda plaques were cored and the inserts excised according to the in vivo excision protocol for Lambda Zap™ II provided by the manufacturer (Stratagene). This in vivo excision protocol moved the cloned insert from the Lambda Hc2 vector into a phagemid vector to allow easy for manipulation and sequencing. The accuracy of the above cloning steps was confirmed by sequencing the insert using the Sanger dideoxy method described in by Sanger et al., *Proc. Natl. Acad. Sci. USA*, 74:5463–5467, (1977) and using the manufacture's instructions in the AMV Reverse Transcriptase ^{35}S-ATP sequencing kit (Stratagene). The sequence of the resulting double-stranded synthetic DNA insert in the V_H expression vector (Lambda Hc2) is shown in FIG. 3. The sequence of each strand (top and bottom) of Lambda Hc2 is listed in the sequence listing as SEQ. ID. NO. 1 and SEQ. ID. NO. 2, respectively. The resultant Lambda Hc2 expression vector is shown in FIG. 4.

(iii) Preparation of Lambda Lc2

To express a plurality of V_L-coding DNA homologs in an *E. coli* host cell, a vector designated Lambda Lc2 was constructed having the capacity to place the V_L-coding DNA homologs in the proper reading frame, provided a ribosome binding site as described by Shine et al., *Nature*, 254:34 (1975), provided the pelB gene leader sequence secretion signal that has been previously used to successfully secrete Fab fragments in *E. coli* by Lei et al., *J. Bac.*, 169:4379 (1987) and Better et al., *Science*, 240:1041 (1988), and also provided a polynucleotide containing a restriction endonuclease site for cloning. Lambda Lc2 has been previously described by Huse et al., *Science*, 246:1275–1281 (1989).

A synthetic DNA sequence containing all of the above features was constructed by designing single stranded polynucleotide segments of 20–60 bases that would hybridize to each other and form the double stranded synthetic DNA sequence shown in FIG. 5. The sequence of each individual single-stranded polynucleotide segment (01–08) within the double stranded synthetic DNA sequence is shown in Table 4.

Polynucleotides 02, 03, 04, 05, 06 and 07 (Table 4) were kinased by adding 1 μl (0.1 ug/μl) of each polynucleotide

and 20 units of T_4 polynucleotide kinase to a solution containing 70 mM Tris-HCl. pH 7.6. 10 mM MgCl. 5 mM DTT. 10 mM beta-mercaptoethanol. 500 mg/ml of BSA. The solution was maintained at 37° C. for 30 minutes and the reaction stopped by maintaining the solution at 65° C. for 10 minutes. The 20 ng each of the two end polynucleotides. 01 and 08. were added to the above kinasing reaction solution together with 1/10 volume of a solution containing 20.0 mM Tris-HCl. pH 7.4. 2.0 mM MgCl and 15.0 mM sodium chloride (NaCl). This solution was heated to 70° C. for 5 minutes and allowed to cool to room temperature. approximately 25° C.. over 1.5 hours in a 500 ml beaker of water. During this time period all 8 polynucleotides annealed to form the double stranded synthetic DNA insert shown in FIG. 5. The individual polynucleotides were covalently linked to each other to stabilize the synthetic DNA insert by adding 40 µl of the above reaction to a solution containing 50 ml Tris-HCl. pH 7.5. 7 ml MgCl. 1 mm DTT. 1 mm ATP and 10 units of T4 DNA ligase. This solution was maintained at 37° C. for 30 minutes and then the T4 DNA ligase was inactivated by maintaining the solution at 65° C. for 10 minutes. The end polynucleotides were kinased by mixing 52 µl of the above reaction. 4 µl of a solution containing 10 mM ATP and 5 units of T4 polynucleotide kinase. This solution was maintained at 37° C. for 30 minutes and then the T4 polynucleotide kinase was inactivated by maintaining the solution at 65° C. for 10 minutes.

TABLE 4

SEQ. ID. NO.		
(32)	01)	5' TGAATTCTAAACTAGTCGCCAAGGAGACAGTCAT 3'
(33)	02)	5' AATGAAATACCTATTGCCTACGGCAGCCGCTGGATT 3'
(34)	03)	5' GTTATTACTCGCTGCCCAACCAGCCATGGCC 3'
(35)	04)	5' GAGCTCGTCAGTTCTAGAGTTAAGCGGCCG 3'
(36)	05)	5' GTATTTCATTATGACTGTCTCCTTGGCGACTAGTTTAGAA-TTCAAGCT 3'
(37)	06)	5' CAGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAG 3'
(38)	07)	5' TGACGAGCTCGGCCATGGCTGGTTGGG 3'
(39)	08)	5' TCGACGGCCGCTTAACTCTAGAAC 3'

The completed synthetic DNA insert was ligated directly into the Lambda Zap™ II vector described in Example 1(a)(i) that had been previously digested with the restriction enzymes Sac I and Xho I. The ligation mixture was packaged according to the manufacture's instructions using Gigapack II Gold packing extract (Stratagene). The packaged ligation mixture was plated on XL1-Blue cells (Stratagene). Individual lambda plaques were cored and the inserts excised according to the in vivo excision protocol for Lambda Zap™ II provided by the manufacturer (Stratagene). This in vivo excision protocol moved the cloned insert from the Lambda Lc2 vector into a plasmid phagemid vector allow for easy manipulation and sequencing. The accuracy of the above cloning steps was confirmed by sequencing the insert using the manufacture's instructions in the AMV Reverse Transcriptase ^{35}S-dATP sequencing kit (Stratagene). The sequence of the resulting Lc2 expression vector (Lambda Lc2) is shown in FIG. 5. Each strand is separately listed in the Sequence Listing as SEQ. ID. NO. 3 and SEQ. ID. NO. 4. The resultant Lc2 vector is schematically diagrammed in FIG. 6.

A preferred vector for use in this invention. designated Lambda Lc3. is a derivative of Lambda Lc2 prepared above. Lambda Lc2 contains a Spe I restriction site (ACTAGT) located 3' to the EcoR I restriction site and 5' to the Shine-Dalgarno ribosome binding site as shown in the sequence in FIG. 5 and in SEQ. ID. NO. 3. A Spe I restriction site is also present in Lambda Hc2 as shown in FIGS. 3 and 4 and in SEQ. ID. NO. 1. A combinatorial vector. designated pComb. was constructed by combining portions of Lambda Hc2 and Lc2 together as described in Example 1a(iv) below. The resultant combinatorial pComb vector contained two Spe I restriction sites. one provided by Lambda Hc2 and one provided by Lambda Lc2. with an EcoR I site in between. Despite the presence of two Spe I restriction sites. DNA homologs having Spe I and EcoR I cohesive termini were successfully directionally ligated into a pComb expression vector previously digested with Spe I and EcoR I as described in Example 1b below. The proximity of the EcoR I restriction site to the 3' Spe I site. provided by the Lc2 vector. inhibited the complete digestion of the 3' Spe I site. Thus. digesting pComb with Spe I and EcoR I did not result in removal of the EcoR I site between the two Spe I sites.

The presence of a second Spe I restriction site may be undesirable for ligations into a pComb vector digested only with Spe I as the region between the two sites would be eliminated. Therefore. a derivative of Lambda Lc2 lacking the second or 3' Spe I site. designated Lambda Lc3. is produced by first digesting Lambda Lc2 with Spe I to form a linearized vector. The ends are filled in to form blunt ends which are ligated together to result in Lambda Lc3 lacking a Spe I site. Lambda Lc3 is a preferred vector for use in constructing a combinatorial vector as described below.

(iv) Preparation of pComb

Phagemids were excised from the expression vectors Lambda Hc2 or Lambda Lc2 using an in vivo excision protocol described above. Double stranded DNA was prepared from the phagemid-containing cells according to the methods described by Holmes et al.. *Anal. Biochem.*, 114:193 (1981). The phagemids resulting from in vivo excision contained the same nucleotide sequences for antibody fragment cloning and expression as did the parent vectors. and are designated phagemid Hc2 and Lc2. corresponding to Lambda Hc2 and Lc2. respectively.

For the construction of combinatorial phagemid vector pComb. produced by combining portions of phagemid Hc2 and phagemid Lc2. phagemid Hc2 was first digested with Sac I to remove the restriction site located 5' to the LacZ promoter. The linearized phagemid was then blunt ended with T4 polymerase and ligated to result in a Hc2 phagemid lacking a Sac I site. The modified Hc2 phagemid and the Lc2 phagemid were then separately restriction digested with Sca I and EcoR I and the linearized phagemids were ligated together at their respective cohesive ends. The ligated phagemid vector was then inserted into an appropriate bacterial host and transformants were selected on the antibiotic ampicillin.

Selected ampicillin resistant transformants were screened for the presence of two Not I sites. The resulting ampicillin resistant combinatorial phagemid vector was designated pComb, the schematic organization of which is shown in FIG. 7. The resultant combinatorial vector, pComb, consisted of a DNA molecule having two cassettes to express two fusion proteins and having nucleotide residue sequences for the following operatively linked elements listed in a 5' to 3' direction: a first cassette consisting of an inducible LacZ promoter upstream from the LacZ gene; a Not I restriction site; a ribosome binding site; a pelB leader; a spacer; a cloning region bordered by a 5' Xho and 3' Spe I restriction site; a decapeptide tag followed by expression control stop sequences; an EcoR I restriction site located 5' to a second cassette consisting of an expression control ribosome binding site; a pelB leader; a spacer region; a cloning region bordered by a 5' Sac I and a 3' Xba I restriction site followed by expression control stop sequences and a second Not I restriction site.

A preferred combinatorial vector for use in this invention, designated pComb2, is constructed by combining portions of phagemid Hc2 and phagemid Lc3 as described above for preparing pComb. The resultant combinatorial vector, pComb2, consists of a DNA molecule having two cassettes identical to pComb to express two fusion proteins identically to pComb except that a second Spe I restriction site in the second cassette is eliminated.

b. Construction of Vectors pCombVIII and pCombIII for Expressing Fusion Proteins Having a Bacteriophage Coat Protein Membrane Anchor

Because of the multiple endonuclease restriction cloning sites, the pComb phagemid expression vector prepared above is a useful cloning vehicle for modification for the preparation of an expression vector of this invention. To that end, pComb is digested with EcoR I and Spe I followed by phosphatase treatment to produce linearized pComb.

(i) Preparation of pCombVIII

A PCR product produced in Example 2g and having a nucleotide sequence that defines a filamentous bacteriophage coat protein VIII (cpVIII) membrane anchor domain and cohesive Spe I and EcoR I termini was admixed with the linearized pComb to form a ligation admixture. The cpVIII-membrane anchor-encoding PCR fragment was directionally ligated into the pComb phagemid expression vector at corresponding cohesive termini, that resulted in forming pCombVIII (also designated pComb8). pCombVIII contains a cassette defined by the nucleotide sequence shown in SEQ. ID. NO. 116 from nucleotide base 1 to base 208, and contains a pelB secretion signal operatively linked to the cpVIII membrane anchor.

A preferred phagemid expression vector for use in this invention, designated either pComb2-VIII or pComb2-8, is prepared as described above by directionally ligating the cpVIII membrane anchor-encoding PCR fragment into a pComb2 phagemid expression vector via Spe I and EcoR I cohesive terminii. The pComb2-8 has only one Spe I restriction site.

(ii) Preparation of pCombIII

A separate phagemid expression vector was constructed using sequences encoding bacteriophage cpIII membrane anchor domain. A PCR product defining the cpIII membrane anchor and Spe I and EcoR I cohesive termini was prepared as described for cpVIII, the details of which are described in Example 2g. The cpIII-derived PCR product was then ligated into linearized pComb vector to form the vector pCombIII (also designated pComb3).

A preferred phagemid expression vector for use in this invention, designated either pComb2-III or pComb2-3, is prepared as described above by directionally ligating the cpIII membrane anchor-encoding PCR fragment into a pComb2 phagemid expression vector via Spe I and Spe I cohesive terminii. The pComb2-3 has only one Spe I restriction site.

c. Construction of pCBAK Vectors Having a Chloramphenicol Resistance Marker

In order to utilize a different selectable marker gene, such as chloramphenicol acetyl transferase (CAT), for the selection of bacteria transformed with a vector of this invention, expression vectors based on pComb were developed having a gene encoding CAT and are designated pCBAK vectors. The pCBAK vectors are prepared by combining portions of pCB and pComb.

(i) Preparation of pCB

pBlueScript phagemid vectors, pBC SK(−) and pBS SK(−). (Stratagene), were modified and combined to generate a third vector designated pCB as described below.

pBC SK(−), which contains a chloramphenicol resistance selectable marker gene, was digested with Bst BI and blunt ended with T4 polymerase. A second digestion with Pvu I allowed for the removal of a 1 kilobase (kb) fragment leaving a 2.4 kb linearized vector which retained the CAT selectable resistance marker gene, an inducible LacZ promoter upstream from the LacZ gene and a ColE1 origin region. The 2.4 kb fragment was recovered. The pBS SK(−) vector was digested with Aat II and blunt ended with T4 polymerase. A second digestion with Pvu I allowed for the isolation of an 800 base pair (bp) fragment containing the f1 origin of replication. Ligation of the pBS derived 800 bp f1 fragment with the 2.4 kb pBC fragment created a pCB precursor vector containing a Sac I site, an f1 origin of replication, a CAT selectable resistance marker gene, ColE1 origin, a multiple cloning site (MCS) flanked by T_3 and T_7 promoters, and an inducible LacZ promoter upstream from LacZ gene.

The pCB precursor vector was then digested with Sac I and blunt-ended with T4 polymerase. The T4 polymerase-treated pCB vector was then religated to form pCB vector and is lacking a Sac I site.

(ii) Preparation of pCBAK0

The pCB vector containing the CAT selectable resistance marker gene was digested with Sac II and Apa I and treated with phosphatase to prevent religation and to form linearized pCB vector. The pComb vector prepared in Example 1(a)(iv) was restriction digested with Sac II and Apa I to release a fragment containing nucleotide residue sequences starting 5' to the LacZ promoter and extending past the 3' end of the second Not I site. The Sac II and Apa I pComb DNA fragment was then directionally ligated into the similarly digested pCB vector to form phagemid expression vector pCBAK0. Preferred pCBAK expression vectors are constructed with pComb2. The resultant pCBAK expression vector contains only one Spe I restriction site.

(iii) Preparation of pCBAK8

To prepare a pCBAK-based phagemid expression vector which encodes a bacteriophage coat protein membrane anchor domain in the expressed fusion protein, pCB phagemid cloning vector prepared in Example 1c(ii) was linearized by digestion with Sac II and Apa I. The pCombVIII phagemid expression vector, prepared in Example 1b(i), was restriction digested with Sac II and Apa I to form a fragment containing a nucleotide residue sequence starting 5' to the LacZ promoter and extending past the 3' end of the

second Not 1 site. The fragment was directionally ligated into the linearized pCB cloning vector to form phagemid expression vector pCBAK8.

(iv) Preparation of pCBAK3

The phagemid expression vector, pCBAK3, for the expression of fusion protein having cpIII membrane anchor domains, was similarly constructed by directionally ligating the Sac II and Apa I restriction digested fragment from pCombIII with Sac II and Apa I linearized pCB cloning vector.

2. Construction of Dicistronic Expression Vectors for Expressing Anti-NPN Heterodimer on Phage Surfaces

In practicing this invention, the heavy (Fd consisting of V_H and C_H1) and light (kappa) chains (V_L, C_L) of antibodies are first targeted to the periplasm of *E. coli* for the assembly of heterodimeric Fab molecules. In order to obtain expression of antibody Fab libraries on a phage surface, the nucleotide residue sequences encoding either the Fd or light chains must be operatively linked to the nucleotide residue sequence encoding a filamentous bacteriophage coat protein membrane anchor. Two preferred coat proteins for use in this invention in providing a membrane anchor are VIII and III (cpVIII and cpIII, respectively). In the Examples described herein, methods for operatively linking a nucleotide residue sequence encoding a Fd chain to either cpVIII or cpIII membrane anchors in a fusion protein of this invention are described.

In a phagemid vector, a first and second cistron consisting of translatable DNA sequences are operatively linked to form a dicistronic DNA molecule. Each cistron in the dicistronic DNA molecule is linked to DNA expression control sequences for the coordinate expression of a fusion protein, Fd-cpVIII or Fd-cpIII, and a kappa light chain.

The first cistron encodes a periplasmic secretion signal (pelB leader) operatively linked to the fusion protein, either Fd-cpVIII or Fd-cpIII. The second cistron encodes a second pelB leader operatively linked to a kappa light chain. The presence of the pelB leader facilitates the coordinated but separate secretion of both the fusion protein and light chain from the bacterial cytoplasm into the periplasmic space.

The process described above is schematically diagrammed in FIG. 8. Briefly, the phagemid expression vector carries a chloramphenicol acetyl transferase (CAT) selectable resistance marker gene in addition to the Fd-cpVIII fusion and the kappa chain. The f1 phage origin of replication facilitates the generation of single stranded phagemid. The isopropyl thiogalactopyranoside (IPTG) induced expression of a dicistronic message encoding the Fd-cpVIII fusion (V_H, C_{H1}, cpVIII) and the light chain (V_L, C_L) leads to the formation of heavy and light chains. Each chain is delivered to the periplasmic space by the pelB leader sequence, which is subsequently cleaved. The heavy chain is anchored in the membrane by the cpVIII membrane anchor domain while the light chain is secreted into the periplasm. The heavy chain in the presence of light chain assembles to form Fab molecules. This same result can be achieved if, in the alternative, the light chain is anchored in the membrane via a light chain fusion protein having a membrane anchor and heavy chain is secreted via a pelB leader into the periplasm.

With subsequent infection of *E. coli* with a helper phage, as the assembly of the filamentous bacteriophage progresses, the coat protein VIII is incorporated along the entire length of the filamentous phage particles as shown in FIGS. 8 and 9. If cpIII is used, the accumulation occurs on the tail of the bacteriophage. The advantage of the utilization of membrane anchors from cpVIII over cpIII is two fold. Firstly, a multiplicity of binding sites, consisting of approximately 2700 cpVIII monomers assembled in a tubular array, exist along the particle surface. Secondly, the construct does not interfere with phage infectivity.

a. Polynucleotide Selection

The nucleotide sequences encoding the immunoglobulin protein CDR's are highly variable. However, there are several regions of conserved sequences that flank the V region domains of either the light or heavy chain, for instance, and that contain substantially conserved nucleotide sequences, i.e., sequences that will hybridize to the same primer sequence. Therefore, polynucleotide synthesis (amplification) primers that hybridize to the conserved sequences and incorporate restriction sites into the DNA homolog produced that are suitable for operatively linking the synthesized DNA fragments to a vector were constructed. More specifically, the primers are designed so that the resulting DNA homologs produced can be inserted into an expression vector of this invention in reading frame with the upstream translatable DNA sequence at the region of the vector containing the directional ligation means.

(i) V_H Primers

For amplification of the V_H domains, primers are designed to introduce cohesive termini compatible with directional ligation into the unique Xho I and Spe I sites of the phagemid Hc2 expression vector. For example, the 3' primer (primer 12A in Table 5), was designed to be complementary to the mRNA in the J_H region. In all cases, the 5' primers (primers 1–10, Table 5) were chosen to be complementary to the first strand cDNA in the conserved N-terminus region (antisense strand). Initially amplification was performed with a mixture of 32 primers (primer 1, Table 5) that were degenerate at five positions. Hybridoma mRNA could be amplified with mixed primers, but initial attempts to amplify mRNA from spleen yielded variable results. Therefore, several alternatives to amplification using the mixed 5' primers were compared.

The first alternative was to construct multiple unique primers, eight of which are shown in Table 5, corresponding to individual members of the mixed primer pool. The individual primers 2–9 of Table 5 were constructed by incorporating either of the two possible nucleotides at three of the five degenerate positions.

The second alternative was to construct a primer containing inosine (primer 10, Table 5) at four of the variable positions based on the published work of Takahashi, et al., *Proc. Natl. Acad. Sci. (U.S.A.)*, 82:1931–1935, (1985) and Ohtsuka et al., *J. Biol. Chem.*, 260:2605–2608, (1985). This primer has the advantage that it is not degenerate and, at the same time minimizes the negative effects of mismatches at the unconserved positions as discussed by Martin et al., *Nuc. Acids Res.*, 13:8927 (1985). However, it was not known if the presence of inosine nucleotides would result in incorporation of unwanted sequences in the cloned V_H regions. Therefore, inosine was not included at the one position that remains in the amplified fragments after the cleavage of the restriction sites. As a result, inosine was not in the cloned insert.

Additional V_H amplification primers including the unique 3' primer were designed to be complementary to a portion of the first constant region domain of the gamma 1 heavy chain mRNA (primers 16 and 17, Table 5). These primers will produce DNA homologs containing polynucleotides coding for amino acids from the V_H and the first constant region

domains of the heavy chain. These DNA homologs can therefore be used to produce Fab fragments rather than an F_V.

Additional unique 3' primers designed to hybridize to similar regions of another class of immunoglobulin heavy chain such as IgM. IgE and IgA are contemplated. Other 3' primers that hybridize to a specific region of a specific class of CH_1 constant region and are adapted for transferring the V_H domains amplified using this primer to an expression vector capable of expressing those V_H domains with a different class of heavy or light chain constant region are also contemplated.

As a control for amplification from spleen or hybridoma mRNA. a set of primers hybridizing to a highly conserved region within the constant region IgG. heavy chain gene were constructed. The 5' primer (primer 11. Table 5) is complementary to the cDNA in the C_H2 region whereas the 3' primer (primer 13. Table 5) is complementary to the mRNA in the C_H3 region. It is believed that no mismatches were present between these primers and their templates.

The primers used for amplification of heavy chain Fd fragments for construction of Fabs are shown at least in Table 5. Amplification was performed in eight separate reactions. each containing one of the 5' primers (primers 2–9) and one of the 3' primers (primer 16). The remaining 5' primers that have been used for amplification in a single reaction are either a degenerate primer (primer 1) or a primer that incorporates inosine at four degenerate positions (primer 10. Table 5. and primers 17 and 18. Table 6). The remaining 3' primer (primer 14. Table 6) has been used to construct F_V fragments. Many of the 5' primers incorporate a Xho I site. and the 3' primers incorporate a Spe I restriction site for insertion of the V_H DNA homolog into the phagemid Hc2 expression vector (FIG. 4).

V_H amplification primers designed to amplify human heavy chain variable regions are shown in Table 6. One of the 5' heavy chain primer contains inosine residues at degenerate nucleotide positions allowing a single primer to hybridize to a large number of variable region sequences. Primers designed to hybridize to the constant region sequences of various IgG mRNAs are also shown in Table 6.

(ii) V_L Primers

The nucleotide sequences encoding the V_L CDRs are highly variable. However. there are several regions of conserved sequences that flank the V_L CDR domains including the J_L. V_L framework regions and V_L leader/promotor. Therefore. amplification primers were constructed that hybridized to the conserved sequences and incorporate restriction sites that allow cloning the amplified fragments into the phagemid Lc2 vector cut with Sac I and Xba I.

For amplification of the V_L CDR domains. the 5' primers (primers 1–8 in Table 6) were designed to be complementary to the first strand cDNA in the conserved N-terminus region. These primers also introduced a Sac I restriction endonuclease site to allow the V_L DNA homolog to be cloned into the phagemid Lc2 expression vector. The 3' V_L amplification primer (primer 9 in Table 6) was designed to be complementary to the mRNA in the J_L regions and to introduce the Xba.I restriction endonuclease site required to insert the V_L DNA homolog into the phagemid Lc2 expression vector (FIG. 6).

Additional 3' V_L amplification primers were designed to hybridize to the constant region of either kappa or lambda mRNA (primers 10 and 11 in Table 6). These primers allow a DNA homolog to be produced containing polynucleotide sequences coding for constant region amino acids of either kappa or lambda chain. These primers make it possible to produce an Fab fragment rather than an F_V.

The primers used for amplification of kappa light chain sequences for construction of Fabs are shown at least in Table 6. Amplification with these primers was performed in 5 separate reactions. each containing one of the 5' primers (primers 3–6. and 12) and one of the 3' primers (primer 13). The remaining 3' primer (primer 9) has been used to construct F_V fragments. The 5' primers contain a Sac I restriction site and the 3' primers contain a Xba I restriction site.

V_L amplification primers designed to amplify human light chain variable regions of both the lambda and kappa isotypes are also shown in Table 6.

All primers and synthetic polynucleotides described herein. including those shown in Tables 3–7 were either purchased from Research Genetics in Huntsville. Ala. or synthesized on an Applied Biosystems DNA synthesizer. model 381A. using the manufacturer's instruction.

TABLE 5

(1)	5'AGGT(C/G)(C/A)A(G/A)CT(G/T)CTCGAGTC(T/A)GG 3'	degenerate 5' primer for the amplification of mouse and human heavy chain variable regions (V_H)
(2)	5'AGGTCCAGCTGCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(3)	5'AGGTCCAGCTGCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(4)	5'AGGTCCAGCTTCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(5)	5'AGGTCCAGCTTCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(6)	5'AGGTCCAACTGCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(7)	5'AGGTCCAACTGCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(8)	5'AGGTCCAACTTCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(9)	5'AGGTCCAACTTCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(10)	5'AGGTNNANCTNCTCGAGTC(T/A)GG 3'	5' degenerate primer containing inosine at 4 degenerate positions for amplification of mouse V_H
(11)	5'GCCCAAGGATGTGCTCACC 3'	5' primer for amplification in the C_H2 region of mouse IgG1

TABLE 5-continued

(12)	5'CTATTAGAATTCAACGGTAACAGTGGTGCCTTGGCCCCA 3'	3' primer for amplification of V_H and introducing a 3' Eco RI site
(12A)	5'CTATTAACTAGTAACGGTAACAGTGGTGCCTTG CCCCA 3'	3' primer for amplification of V_H using 3' Spe I site
(13)	5'CTCAGTATGGTGGTTGTGC 3'	3' primer for amplification in the C_H3 region of mouse IgG1
(14)	5'GCTACTAGTTTTGATTTCCACCTTGG 3'	3' primer for amplification of mouse kappa light chain variable regions (V_L)
(15)	5'CAGCCATGGCCGACATCCAGATG 3'	5' primer for amplification of mouse kappa light chain variable regions
(16)	5'AATTTTACTAGTCACCTTGGTGCTGCTGGC 3'	Unique 3' primer for amplification of V_H including part of the mouse gamma 1 first constant region
(17)	5'TATGCAACTAGTACAACCACAATCCCTGGGCACAATTTT 3'	Unique 3' primer for amplification of Fd including part of mouse IgG1 first constant region and hinge region
(18)	5'AGGCTTACTAGTACAATCCCTGGGCACAAT 3'	3' primer for amplifying mouse Fd including part of the mouse IgG first constant region and part of the hinge region

TABLE 6

(1)	5' CCAGTTCCGAGCTCGTTGTGACTCAGGAATCT 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(2)	5' CCAGTTCCGAGCTCGTGTTGACGCAGCCGCCC 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(3)	5' CCAGTTCCGAGCTCGTGCTCACCCAGTCTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(4)	5' CCAGTTCCGAGCTCCAGATGACCCAGTCTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(5)	5' CCAGATGTGAGCTCGTGATGACCCAGACTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(6)	5' CCAGATGTGAGCTCGTCATGACCCAGTCTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(7)	5' CCAGATGTGAGCTCTTGATGACCCAAACTCAA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(8)	5' CCAGATGTGAGCTCGTGATAACCCAGGATGAA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(9)	5' GCAGCATTCTAGAGTTTCAGCTCCAGCTTGCC 3'	Unique 3' primer for amplification of kappa light chain variable regions
(10)	5' CCGCCGTCTAGAACACTCATTCCTGTTGAAGCT 3'	Unique 3' primer for mouse kappa light chain amplification including the constant region
(11)	5' CCGCCGTCTAGAACATTCTGCAGGAGACAGACT 3'	Unique 3' primer for mouse lambda light chain amplification including the constant region
(12)	5' CCAGTTCCGAGCTCGTGATGACACAGTCTCCA 3'	Unique 5' primer for V_L amplification
(13)	5' GCGCCGTCTAGAATTAACACTCATTCCTGTTGAA 3'	Unique 3' primer for amplification of kappa light chain
(14)	5' CTATTAACTAGTAACGGTAACAGTGGTGCCTTGCCCCA 3'	Unique 3' primer for amplification of mouse F_V
(15)	5' AGGCTTACTAGTACAATCCCTGGGCACAAT 3'	Unique 3' primer for amplification of mouse IgG Fd
(16)	5' GCCGCTCTAGAACACTCATTCCTGTTGAA 3'	Unique 3' primer for amplification of mouse kappa light chain
(17)	5' AGGTIIAICTICTCGAGTCTGC 3'	Degenerate 5' primer containing inosine at 4 degenerate positions for amplifying mouse V_H
(18)	5' AGGTIIAICTICTCGAGTCAGC 3'	Degenerate 5' primer containing inosine at 4 degenerate positions for amplifying mouse V_H
(19)	5' GTGCCAGATGTGAGCTCGTGATGACCCAGTCTCCA 3'	Unique 5' primer for human and mouse kappa V_L amplification
(20)	5' TCCTTCTAGATTACTAACACTCTCCCCTGTTGAA 3'	Unique 3' primer for kappa V_L amplification
(21)	5' GCATTCTAGACTATTATGAACATTCTGTAGGGGC 3'	Unique 3' primer for human, mouse and rabbit lambda V_L amplification
(22)	5' CTGCACAGGGTCCTGGGCCGAGCTCGTGGTGACTCAG 3'	Unique 5' primer for human lambda V_L amplification
(23)	5' AGTTGCAIITGCTCGAGTCTGG 3'	5' degenerate primer for human V_H amplification containing inosine at 3 degenerate positions
(24)	5' GTGGGCATGTGTGAGTTGTGTCACTAGTTGGGGTTTTGAGCTC 3'	Unique 3' primer for human V_H amplification
(25)	5' AGCATCACTAGTACAAGATTTGGGCTC 3'	Unique 3' primer for human IgG1 Fd

TABLE 6-continued

		amplification
(26)	5' AGCATCACTAGTACAAGATTTGGGCTC 3'	Unique 3' primer for amplification of human variable regions (V_H)
(27)	5' AGGTGCAGCTGCTCGAGTCGGG 3'	Unique 3' primer for amplification of human variable regions (V_H)
(28)	5' AGGTGCAACTGCTCGAGTCTGG 3'	Unique 3' primer for amplification of human variable regions (V_H)
(29)	5' AGGTGCAACTGCTCGAGTCGGG 3'	Unique 3' primer for amplification of human variable regions (V_H)
(30)	5' TCCTTCTAGATTACTAACACTCTCCCCTGTTGAA 3'	3' primer in human kappa light chain constant region
(31)	5' CTGCACAGGGTCCTGGGCCGAGCTCGTGGTGACTCAG 3'	5' primer for amplification of human lambda light chain variable regions
(32)	5' GCATTCTAGACTATTAACATTCTGTAGGGGC 3'	3' primer in human lambda light chain constant region
(33)	5' ACCCAAGGACACCCTCATG 3'	Control primer hybridizing to the human CH_2 region
(34)	5' CTCAGTATGGTGGTTGTGC 3'	Control primer hybridizing to the human CH_3 region
(35)	5' GTCTCACTAGTCTCCACCAAGGGCCCATCGGTC 3'	5' primer for amplifying human IgG heavy chain first constant region
(36)	5' ATATACTAGTGAGACAGTGACCAGGGTTCCTTGGCCCCA 3'	3' primer for amplifying human heavy chain variable regions
(37)	5' ACGTCTAGATTCCACCTTGGTCCC 3'	3' primer for amplifying human kapp chain variable regions
(38)	5' GCATACTAGTCTATTAACATTCTGTAGGGGC 3'	5' primer for amplifying human kappa light chain constant region
(39)	5' CCGGAATTCTTATCATTTACCCGGAGA 3'	3' primer located in the CH3 region of human IgG1 to amplify the entire heavy chain
(40)	5' TCTGCACTAGTTGGAATGGGCACATGCAG 3'	3' primer for amplifying the Fd region of mouse IgM

The 19 primers listed in Table 5 have been listed in the Sequence Listing and have been assigned the following SEQ. ID. NO.:

(1)=SEQ. ID. NO. 40
(2)=SEQ. ID. NO. 41
(3)=SEQ. ID. NO. 42
(4)=SEQ. ID. NO. 43
(5)=SEQ. ID. NO. 44
(6)=SEQ. ID. NO. 45
(7)=SEQ. ID. NO. 46
(8)=SEQ. ID. NO. 47
(9)=SEQ. ID. NO. 48
(10)=SEQ. ID. NO. 49
(11)=SEQ. ID. NO. 50
(12)=SEQ. ID. NO. 51
(12A)=SEQ. ID. NO. 52
(13)=SEQ. ID. NO. 53
(14)=SEQ. ID. NO. 54
(15)=SEQ. ID. NO. 55
(16)=SEQ. ID. NO. 56
(17)=SEQ. ID. NO. 57
(18)=SEQ. ID. NO. 58

The 40 primers listed as "(1)" through "(40)" in Table 6 have also been individually and sequentially listed in the Sequence Listing beginning with SEQ. ID. NO. 59 through SEQ. ID. NO. 98, respectively.

b. Preparation of a Repertoire of Genes Encoding Immunoglobulin Variable Domain

Nitrophenylphosphonamidate (NPN) was selected as the ligand for receptor binding in preparing a heterodimeric receptor according to the methods of the invention.

Keyhole limpet hemocyanin (KLH) was conjugated to NPN to form a NPN-KLH conjugate used for immunizing a mouse to produce an anti-NPN immune response and thereby provide a source of ligand specific heterodimeric receptor genes.

The NPN-KLH conjugate was prepared by admixing 250 μl of a solution containing 2.5 mg of NPN in dimethylformamide with 750 μl of a solution containing 2 mg of KLH in 0.01 Molar (M) sodium phosphate buffer (pH 7.2). The two solutions were admixed by slow addition of the NPN solution to the KLH solution while the KLH solution was being agitated by a rotating stirring bar. Thereafter the admixture was maintained at 4° C. for 1 hour with the same agitation to allow conjugation to proceed. The conjugated NPN-KLH was isolated from the nonconjugated NPN and KLH by gel filtration through Sephadex G-25. The isolated NPN-KLH conjugate was injected into mice as described below.

The NPN-KLH conjugate was prepared for injection into mice by adding 100 μg of the conjugate to 250 μl of phosphate buffered saline (PBS). An equal volume of complete Freund's adjuvant was added and emulsified the entire solution for 5 minutes. A 129 G_{IX+} mouse was injected with 300 μl of the emulsion. Injections were given subcutaneously at several sites using a 21 gauge needle. A second immunization with NPN-KLH was given two weeks later. This injection was prepared as follows: 50 micrograms (μg) of NPN-KLH were diluted in 250 μl of PBS and an equal volume of alum was admixed to the NPN-KLH solution. The mouse was injected intraperitoneally with 500 μl of the solution using a 23 gauge needle. One month later the mice were given a final injection of 50 μg of the NPN-KLH conjugate diluted to 200 μl in PBS. This injection was given intravenously in the lateral tail vein using a 30 gauge needle. Five days after this final injection the mice were sacrificed and total cellular RNA was isolated from their spleens.

Total cellular RNA was prepared from the spleen of a single mouse immunized with KLH-NPN as described above using the RNA preparation methods described by Chomczynski et al., *Anal Biochem.*, 162:156–159 (1987) and using the RNA isolation kit (Stratagene) according to the manufacturer's instructions. Briefly, immediately after

removing the spleen from the immunized mouse. the tissue was homogenized in 10 ml of a denaturing solution containing 4.0M guanine isothiocyanate. 0.25M sodium citrate at pH 7.0. and 0.1M beta-mercaptoethanol using a glass homogenizer. One ml of sodium acetate at a concentration of 2M at pH 4.0 was admixed with the homogenized spleen. One ml of phenol that had been previously saturated with H_2O was also admixed to the denaturing solution containing the homogenized spleen. Two ml of a chloroform:isoamyl alcohol (24:1 v/v) mixture was added to this homogenate. The homogenate was mixed vigorously for ten seconds and maintained on ice for 15 minutes. The homogenate was then transferred to a thick-walled 50 ml polypropylene centrifuged tube (Fisher Scientific Company. Pittsburg. Pa.). The solution was centrifuged at 10.000×g for 20 minutes at 4° C. The upper RNA-containing aqueous layer was transferred to a fresh 50 ml polypropylene centrifuge tube and mixed with an equal volume of isopropyl alcohol. This solution was maintained at −20° C. for at least one hour to precipitate the RNA. The solution containing the precipitated RNA was centrifuged at 10.000×g for twenty minutes at 4° C. The pelleted total cellular RNA was collected and dissolved in 3 ml of the denaturing solution described above. Three ml of isopropyl alcohol was added to the re-suspended total cellular RNA and vigorously mixed. This solution was maintained at −20° C. for at least 1 hour to precipitate the RNA. The solution containing the precipitated RNA was centrifuged at 10.000×g for ten minutes at 4° C. The pelleted RNA was washed once with a solution containing 75% ethanol. The pelleted RNA was dried under vacuum for 15 minutes and then re-suspended in dimethyl pyrocarbonate (DEPC) treated (DEPC-H_2O) H_2O.

Messenger RNA (mRNA) enriched for sequences containing long poly A tracts was prepared from the total cellular RNA using methods described in *Molecular Cloning: A Laboratory Manual*. Maniatis et al., eds.. Cold Spring Harbor. N.Y.. (1982). Briefly. one half of the total RNA isolated from a single immunized mouse spleen prepared as described above was re-suspended in one ml of DEPC-H_2O and maintained at 65° C. for five minutes. One ml of 2×high salt loading buffer consisting of 100 mM Tris-HCl (Tris [hydroxymethyl] amino methane hydrochloride). 1M sodium chloride (NaCl). 2.0 mM disodium ethylene diamine tetra-acetic acid (EDTA) at pH 7.5. and 0.2% sodium dodecyl sulfate (SDS) was added to the re-suspended RNA and the mixture allowed to cool to room temperature. The mixture was then applied to an oligo-dT (Collaborative Research Type 2 or Type 3) column that was previously prepared by washing the oligo-dT with a solution containing 0.1M sodium hydroxide and 5 mM EDTA and then equilibrating the column with DEPC-H_2O. The eluate was collected in a sterile polypropylene tube and reapplied to the same column after heating the eluate for 5 minutes at 65° C. The oligo dT column was then washed with 2 ml of high salt loading buffer consisting of 50 mM Tris-HCl. pH 7.5. 500 mM sodium chloride. 1 mM EDTA at pH 7.5 and 0.1% SDS. The oligo dT column was then washed with 2 ml of 1×medium salt buffer consisting of 50 mM Tris-HCl. pH 7.5. 100 mM. 1 mM EDTA and 0.1% SDS. The messenger RNA was eluted from the oligo dT column with 1 ml of buffer consisting of 10 mM Tris-HCl. pH 7.5. 1 mM EDTA. at pH 7.5. and 0.05% SDS. The messenger RNA was purified by extracting this solution with phenol/chloroform followed by a single extraction with 100% chloroform. The messenger RNA was concentrated by ethanol precipitation and re-suspended in DEPC H_2O.

The messenger RNA (mRNA) isolated by the above process contains a plurality of different V_H coding polynucleotides. i.e.. greater than about 10^4 different V_H-coding genes. and contains a similar number of V_L-coding genes. Thus. the mRNA population represents a repertoire of variable region-coding genes.

c. Preparation of DNA Homologs

In preparation for PCR amplification. mRNA prepared above is used as a template for cDNA synthesis by a primer extension reaction. In a typical 50 μl transcription reaction. 5–10 ug of spleen mRNA in water is first hybridized (annealed) with 500 ng (50.0 pmol) of the 3' V_H primer (primer 12A. Table 5). at 65° C. for five minutes. Subsequently. the mixture is adjusted to 1.5 mM dATP. dCTP. dGTP and dTTP. 40 mM Tris-HCl. pH 8.0. 8 mM $MgCl_2$. 50 mM NaCl. and 2 mM spermidine. Moloney-Murine Leukemia virus Reverse transcriptase (Stratagene). 26 units. is added and the solution is maintained for 1 hour at 37 C.

PCR amplification is performed in a 100 μl reaction containing the products of the reverse transcription reaction (approximately 5 ug of the cDNA/RNA hybrid). 300 ng of 3' V_H primer (primer 12A of Table 5). 300 ng each of the 5' V_H primers (primers 2–10 of Table 5) 200 mM of a mixture of dNTP's. 50 mM KCl. 10 mM Tris-HCl pH 8.3. 15 mM $MgCl_2$. 0.1% gelatin and 2 units of *Thermus aquaticus* (Taq) DNA polymerase. The reaction mixture is overlaid with mineral oil and subjected to 40 cycles of amplification. Each amplification cycle includes denaturation at 92° C. for 1 minute. annealing at 52° C. for 2 minutes and polynucleotide synthesis by primer extension (elongation) at 72° C. for 1.5 minutes. The amplified V_H-coding DNA homolog containing samples are then extracted twice with phenol/chloroform. once with chloroform. ethanol precipitated and are stored at −70° C. in 10 mM Tris-HCl. pH 7.5. and 1 mM EDTA.

Using unique 5' primers (2–9. Table 5). efficient V_H-coding DNA homolog synthesis and amplification from the spleen mRNA is achieved as shown by agarose gel electrophoresis. The amplified cDNA (V_H-coding DNA homolog) was seen as a major band of the expected size (360 bp). The amount the amplified V_H-coding polynucleotide fragment in each reaction is similar. indicating that all of these primers were about equally efficient in initiating amplification. The yield and quality of the amplification with these primers is reproducible.

The primer containing inosine also synthesizes amplified V_H-coding DNA homologs from spleen mRNA reproducibly. leading to the production of the expected sized fragment. of an intensity similar to that of the other amplified cDNAs. The presence of inosine also permits efficient DNA homolog synthesis and amplification. clearly indicating that such primers are useful in generating a plurality of V_H-coding DNA homologs. Amplification products obtained from the constant region primers (primers 11 and 13. Table 5) are more intense indicating that amplification was more efficient. possibly because of a higher degree of homology between the template and primers. Following the above procedures. a V_H-coding gene library is constructed from the products of eight amplifications. each performed with a different 5' primer. Equal portions of the products from each primer extension reaction are mixed and the mixed product is then used to generate a library of V_H-coding DNA homolog-containing vectors.

DNA homologs of the V_L are also prepared from the purified mRNA prepared as described above. In preparation for PCR amplification. mRNA prepared according to the above examples is used as a template for cDNA synthesis. In a typical 50 μl transcription reaction. 5–10 ug of spleen

mRNA in water is first annealed with 300 ng (50.0 pmol) of the 3' V_L primer (primer 14, Table 5), at 65° C. for five minutes. Subsequently, the mixture is adjusted to 1.5 mM dATP, dCTP, dGTP, and dTTP, 40 mM Tris-HCl, pH 8.0, 8 mM $MgCl_2$, 50 mM NaCl, and 2 mM spermidine. Moloney-Murine Leukemia virus reverse transcriptase (Stratagene), 26 units, is added and the solution is maintained for 1 hour at 37° C. The PCR amplification is performed in a 100 µl reaction containing approximately 5 ug of the cDNA/RNA hybrid produced as described above, 300 ng of the 3' V_L primer (primer 14 of Table 5), 300 ng of the 5' V_L primer (primer 16 of Table 5), 200 mM of a mixture of dNTP's, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 15 mM $MgCl_2$, 0.1% gelatin and 2 units of Taq DNA polymerase. The reaction mixture is overlaid with mineral oil and subjected to 40 cycles of amplification. Each amplification cycle includes denaturation at 92° C. for 1 minute, annealing at 52° C. for 2 minutes and elongation at 72° C. for 1.5 minutes. The amplified samples are then extracted twice with phenol/chloroform, once with chloroform, ethanol precipitated and are stored at −70° C. in 10 mM Tris-HCl, 7.5 and 1 mM EDTA.

d. Insertion of DNA Homologs into a DNA Expression Vector

To prepare an expression library enriched in V_H sequences, DNA homologs enriched in V_H sequences are prepared according to Example 2c using the same set of 5' primers but with primer 12A (Table 5) as the 3' primer. The resulting PCR amplified products (2.5 ug/30 µl of 150 mM NaCl, 8 mM Tris-HCl, pH 7.5, 6 mM $MgSO_4$, 1 mM DTT, 200 ug/ml BSA) are digested at 37° C. with restriction enzymes Xho I (125 units) and Spe I (125 units). In cloning experiments which required a mixture of the products of the amplification reactions, equal volumes (50 µl, 1–10 ug concentration) of each reaction mixture are combined after amplification but before restriction digestion. The V_H homologs are purified on a 1% agarose gel using the standard electro-elution technique described in *Molecular Cloning A Laboratory Manual*, Maniatis et al., eds., Cold Spring Harbor, N.Y., (1982). After gel electrophoresis of the digested PCR amplified spleen mRNA, the region of the gel containing DNA fragments of approximate 350 bps is excised, electro-eluted into a dialysis membrane, ethanol precipitated and re-suspended suspended in a TE solution containing 10 mM Tris-HCl, pH 7.5 and 1 mM EDTA to a final concentration of 50 ng/µl. The resulting V_H DNA homologs represent a repertoire of polypeptide genes having cohesive termini adapted for directional ligation to the vector Lambda Hc2. These prepared V_H DNA homologs are then directly inserted by directional ligation into linearized Lambda Hc2 expression vector prepared as described below.

The Lambda Hc2 expression DNA vector is prepared for inserting a DNA homolog by admixing 100 ug of this DNA to a solution containing 250 units each of the restriction endonucleases Xho I and Spe I (both from Boehringer Mannheim, Indianapolis, Ind.) and a buffer recommended by the manufacturer. This solution is maintained at 37 from 1.5 hours. The solution is heated at 65° C. for 15 minutes top inactivate the restriction endonucleases. The solution is chilled to 30° C. and 25 units of heat-killable (HK) phosphatase (Epicenter, Madison, Wis.) and $CaCl_2$ is admixed to it according to the manufacturer's specifications. This solution is maintained at 30° C. for 1 hour. The DNA is purified by extracting the solution with a mixture of phenol and chloroform followed by ethanol precipitation. The Lambda Hc2 expression vector is now ready for ligation to the V_H DNA homologs prepared in the above examples. These prepared V_H DNA homologs are then directly inserted into the Xho I and Spe I restriction digested Lambda Hc2 expression vector that prepared above by ligating 3 moles of V_H DNA homolog inserts with each mole of the Hc2 expression vector overnight at 5° C. Approximately 3.0×10^5 plague forming units are obtained after packaging the DNA with Gigapack II Bold (Stratagene) of which 50% are recombinants. The ligation mixture containing the V_H DNA homologs are packaged according to the manufacturers specifications using Gigapack Gold II Packing Extract (Stratagene). The resulting Lambda Hc2 expression libraries are then transformed into XL1-Blue cells.

To prepare a library enriched in V_L sequences, PCR amplified products enriched in V_L sequences are prepared according to Example 2c. These V_L DNA homologs are digested with restriction enzymes Sac I and Xba I and the digested V_L DNA homologs are purified on a 1% agarose gel as described above for the V_H DNA homologs to form a repertoire of V_L-polypeptide genes adapted for directional ligation. The prepared V_L DNA homologs are then directionally ligated into the Lambda Lc2 expression vector previously digested with the restriction enzymes, Sac I and Xba I as described for Lambda Hc2. The ligation mixture containing the V_L DNA homologs is packaged to form a Lambda Lc2 expression library as described above and is ready to be plated on XL1-Blue cells.

e. Randomly Combining V_H and V_L DNA Homologs on the Same Expression Vector

The construction of a library containing vectors for expressing two cistrons that express heavy and light chains is accomplished in two steps. In the first step, separate heavy and light chain libraries are constructed in the expression vectors Lambda Hc2 and Lambda Lc2, respectively, as described using gene repertoires obtained from a mouse immunized with NPN-KLH. In the second step, these two libraries are combined at the antisymmetric EcoR I sites present in each vector. This resulted in a library of clones each of which potentially co-expresses a heavy and a light chain. The actual combinations are random and do not necessarily reflect the combinations present in the B-cell population in the parent animal.

The spleen mRNA resulting from the above immunizations (Example 2b) is isolated and used to create a primary library of V_H gene sequences using the Lambda Hc2 expression vector. The primary library contains 1.3×10^6 plaque-forming units (pfu) and can be screened for the expression of the decapeptide tag to determine the percentage of clones expressing V_H and C_H1 (Fd) sequences. The sequence for this peptide is only in frame for expression following the cloning of a Fd (or V_H) fragment into the vector. At least 80% of the clones in the library express Fd fragments based on immunodetection of the decapeptide tag.

The light chain library is constructed in the same way as the heavy chain and contains 2.5×10^6 members. Plaque screening, using an anti-kappa chain antibody, indicates that 60% of the library contained express light chain inserts. A small percentage of inserts results from incomplete dephosphorylation of vector after cleavage with Sac I and Xba I.

Once obtained, the two libraries are used to construct a combinatorial library by crossing them at the EcoR I site. To accomplish the cross, DNA is first purified from each library.

The Lambda Lc2 library prepared in Example 2d is amplified and 500 ug of Lambda Lc2 expression library phage DNA is prepared from the amplified phage stock using the procedures described in *Molecular Cloning: A Laboratory Manual*, Maniatis et al., eds., Cold Spring Harbor, N.Y. (1982). Fifty ug of this amplified expression

library phage DNA is maintained in a solution containing 100 units of MLu I restriction endonuclease (Boehringer Mannheim, Indianapolis, Ind.) in 200 µl of a buffer supplied by the endonuclease manufacturer for 1.5 hours at 37° C. The solution is then extracted with a mixture of phenol and chloroform. The DNA is then ethanol precipitated and re-suspended in 100 µl of water. This solution is admixed with 100 units of the restriction endonuclease EcoR I (Boehringer) in a final volume of 200 µl of buffer containing the components specified by the manufacturer. This solution is maintained at 37° C. for 1.5 hours and the solution is then extracted with a mixture of phenol and chloroform. The DNA was ethanol precipitated and the DNA re-suspended in TE.

The Lambda Hc2 expression library prepared in Example 2d is amplified and 500 ug of Lambda Hc2 expression library phage DNA is prepared using the methods detailed above. 50 ug of this amplified library phage DNA is maintained in a solution containing 100 units of Hind III restriction endonuclease (Boehringer) in 200 µl of a buffer supplied by the endonuclease manufacturer for 1.5 hours at 37° C. The solution is then extracted with a mixture of phenol and chloroform saturated with 0.1M Tris-HCl, pH 7.5. The DNA is then ethanol precipitated and re-suspended in 100 µl of water. This solution is admixed with 100 units of the restriction endonuclease EcoR I (Boehringer) in a final volume of 200 µl of buffer containing the components specified by the manufacturer. This solution is maintained at 37° C. for 1.5 hours and the solution is then extracted with a mixture of phenol and chloroform. The DNA is ethanol precipitated and the DNA re-suspended in TE.

The restriction digested Hc2 and Lc2 expression libraries are ligated together. To that end, a DNA admixture consists of 1 ug of Hc2 and 1 ug of Lc2 phage library DNA is prepared in a 10 µl reaction using the reagents supplied in a ligation kit (Stratagene). The DNA admixture is warmed to 45° C. for 5 minutes to melt any cohesive termini that may have reannealed. The admixture is then chilled to 0° C. to prevent religation. Bacteriophage T4 DNA ligase (0.1 Weiss units which is equivalent to 0.02 units as determined in an exonuclease resistance assay) is admixed into the chilled DNA solution along with 1 ul of 5 mM ATP and 1 ul 10×bacteriophage T4 DNA ligase buffer (10×buffer is prepared by admixing 200 mM Tris-HCl, pH 7.6, 50 mM $Mgcl_2$, 50 mM DTT, and 500 ug/ml BSA) to form a ligation admixture. After ligation for 16 hr at 4° C., 1 µl of the ligated the phage DNA is packaged with Gigapack Gold II packaging extract and plated on XL1-Blue cells prepared according to the manufacturers instructions to form a Lambda phage library of dicistronic expression vectors capable of expressing heavy and light chains derived from the NPN-immunized mouse. A portion of the clones obtained are used to determine the effectiveness of the combination.

f. Selection of Anti-NPN Reactive Heterodimer-Producing Dicistronic Vectors

The combinatorial Fab expression library prepared above in Example 2a was screened to identify clones having affinity for NPN. To determine the frequency of the phage clones which co-expressed the light and heavy chain fragments, duplicate lifts of the light chain, heavy chain and combinatorial libraries were screened as above for light and heavy chain expression. In this study of approximately 500 recombinant phage, approximately 60% co-expressed light and heavy chain proteins.

All three libraries, the light chain, the heavy chain and the combinatorial, were screened to determine if they contained recombinant phage that expressed antibody fragments which bound NPN. In a typical procedure 30,000 phage were plated on XL1-Blue cells and duplicate lifts with nitrocellulose were screened for binding to NPN coupled to ^{125}I labeled BSA. The BSA was iodinated following the Chloramine-T method as described by Bolton et al., *Biochem.*, 133:529–534 (1973). Duplicate screens of 80,000 recombinant phage from the light chain library and a similar number from the heavy chain library did not identify any clones which bound the antigen. In contrast, the screen of a similar number of clones from the Fab expression library identified many phage plaques that bound NPN. This observation indicates that under conditions where many heavy chains in combination with light chains bind to antigen the same heavy or light chains alone do not. Therefore, in the case of NPN, it is believed that there are many heavy and light chains that only bind antigen when they are combined with specific light and heavy chains respectively.

To assess the ability to screen large numbers of clones and obtain a more quantitative estimate of the frequency of antigen binding clones in the combinatorial library, one million phage plaques were screened and approximately 100 clones which bound to antigen were identified. For six clones which were believed to bind NPN, a region of the plate containing the six positive and approximately 20 surrounding bacteriophage plaques was selected and each plaque was cored, replated, and screened with duplicate lifts. As expected, approximately one in twenty of the phage specifically bound to antigen. Cores of regions of the plated phage believed to be negative did not give positives on replating.

Clone 2b, one of the plaques which reacted with NPN, was excised according to an in vivo excision protocol where 200 ul of phage stock and 200 ul of a F+ derivative of XL1-Blue (A_{600}=1.00) (Stratagene) were admixed with 1 ul of M13mp8 helper phage (1×10^{10} pfu/milliliter (ml)) and maintained at 37° C. for 15 minutes. After a four hour maintenance in Luria-Bertani (LB) medium and heating at 70° C. for 20 minutes to heat kill the XL1-Blue cells, the phagemids were re-infected into XL1-Blue cells and plated onto LB plates containing ampicillin. This procedure converted the cloned insert from the Lambda Zap II vector into a plasmid vector to allow easy manipulation and sequencing (Stratagene). The phagemid DNA encoding the V_H and part of the V_L was then determined by DNA sequencing using the Sanger dideoxy method described in Sanger et al., *Proc. Natl. Acad. Sci.*, 74:5463–5467 (1977) using a Sequenase kit according to manufacturer's instructions (US Biochemical Corp., Cleveland, Ohio). The nucleotide residue sequence of Clone 2b Fd chain is listed in the Sequence Listing as SEQ. ID. NO. 99. The nucleotide residue sequences of the kappa light chain variable and constant regions are listed in the Sequence Listing as SEQ. ID. NO. 100 and SEQ. ID. NO. 101, respectively.

g. Preparation of a DNA Sequence Encoding a Filamentous Phage Coat Protein Membrane Anchor

cpVIII Membrane Anchor: M13mp18, a commercially available bacteriophage vector (Pharmacia, Piscataway, N.J.), was used as a source for isolating the gene encoding cpVIII. The sequence of the gene encoding the membrane anchor domain of cpVIII listed in Sequence Listing as SEQ. ID. NO. 102, was modified through PCR amplification to incorporate the restriction endonuclease sites, Spe I and EcoR I, and two stop codons prior to the EcoR I site. The corresponding amino acid residue sequence of the membrane anchor domain of cpVIII is listed as SEQ. ID. NO. 17.

To prepare a modified cpVIII, replicative form DNA from M13mp18 was first isolated. Briefly, into 2 ml of LB

(Luria-Bertani medium). 50 ul of a culture of a bacterial strain carrying an F' episome (JM107, JM109 or TG1) was admixed with a one tenth suspension of bacteriophage particles derived from a single plaque. The admixture was incubated for 4 to 5 hours at 37° C. with constant agitation. The admixture was then centrifuged at 12,000×g for 5 minutes to pellet the infected bacteria. After the supernatant was removed, the pellet was resuspended by vigorous vortexing in 100 ul of ice-cold solution I. Solution I was prepared by admixing 50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl, pH 8.0, and autoclaving for 15 minutes.

To the bacterial suspension, 200 ul of freshly prepared Solution II was admixed and the tube was rapidly inverted five times. Solution II was prepared by admixing 0.2N NaOH and 1% SDS. To the bacterial suspension, 150 ul of ice-cold Solution III was admixed and the tube was vortexed gently in an inverted position for 10 seconds for to disperse Solution III through the viscous bacterial lysate. Solution III was prepared by admixing 60 ml of 5M potassium acetate, 11.5 ml of glacial acetic acid and 28.5 ml of water. The resultant bacterial lysate was then stored on ice for 5 minutes followed by centrifugation at 12,000×g for 5 minutes at 4° C. in a microfuge. The resultant supernatant was recovered and transferred to a new tube. To the supernatant was added an equal volume of phenol: chloroform and the admixture was vortexed. The admixture was then centrifuged at 12,000×g for 2 minutes in a microfuge. The resultant supernatant was transferred to a new tube and the double-stranded bacteriophage DNA was precipitated with 2 volumes of ethanol at room temperature. After allowing the admixture to stand at room temperature for 2 minutes, the admixture was centrifuged to pellet the DNA. The supernatant was removed and the pelleted replicative form DNA was resuspended in 25 ul of Tris-HCl, pH 7.6, and 10 mM EDTA (TE).

The double-stranded M13mp18 replicative form DNA was then used as a template for PCR. Primers, AK 5 (SEQ. ID. NO. 103) and AK 6 (SEQ. ID. NO. 104), the sequences of which are listed in Table 7 below, were used in the PCR reaction to amplify the mature gene for cpVIII member anchor domain and incorporate the two cloning sites, Spe I and EcoR I. For the PCR reaction, 2 ul containing 1 nanogram (ng) of M13mp18 replicative form DNA was admixed with 10 ul of 10×PCR buffer purchased commercially (Promega Biotech, Madison, Wis.) in a 0.5 ml microfuge tube. To the DNA admixture, 8 ul of a 2.5 mM solution of dNTPs (dATP, dCTP, dGTP, dTTP) was admixed to result in a final concentration of 200 micromolar (uM). Three ul (equivalent to 60 picomoles (pM)) of the 5' forward AK 5 primer and 3 ul (60 pM) of the 3' backward AK 6 primer was admixed into the DNA solution. To the admixture, 73 ul of sterile water and 1 ul/5 units of polymerase (Promega Biotech) was added. Two drops of mineral oil were placed on top of the admixture and 40 rounds of PCR amplification in a thermocycler were performed. The amplification cycle consisted of 52° C. for 2 minutes, 72° C. for 1.5 minutes and 91° C. for 2 minutes. The resultant PCR modified cpVIII membrane anchor domain DNA fragment from M13mp18 containing samples were then purified with Gene Clean (BIO101, La Jolla, Calif.), extracted twice with phenol/chloroform, once with chloroform followed by ethanol precipitation and were stored at −70° C. in 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA.

TABLE 7

SEQ. ID NO.	Primer	
(103)[1]	AK 5 (F)	5' GTGCCCAGGGATTGTACTAGTGCTGAGGGTGACGAT 3'
(104)[2]	AK 6 (B)	5' ACTCGAATTCTATCAGCTTGCTTTCGAGGTGAA 3'
(105)[3]	Hc3 (F)	5' AGGTCCAGCTTCTCGAGTCTGG 3'
(106)[4]	AK 7 (B)	5' GTCACCCTCAGCACTAGTACAATCCCTGGGCAC 3'
(107)[5]	G-3 (F)	5' GAGACGACTAGTGGTGGCGGTGGCTCTCCATTC GTTTGTGAATATCAA 3'
(108)[6]	G-3 (B)	5' TTACTAGCTAGCATAATAACGGAATACCCAAAA GAACTGG 3'
(109)[7]	LAC-F	5' TATGCTAGCTAGTAACACGACAGGTTTCCCGAC TGG 3'
(110)[8]	LAC-B	5' ACCGAGCTCGAATTCGTAATCATGGTC 3'

F Forward Primer
B Backward Primer
[1]From 5' to 3': the overlapping sequence for C_H1 3' end is double underlined; the Spe I restriction site sequence is single underlined; the overlapping sequence for cpVIII is double underlined.
[2]EcoR I restriction site sequence is single underlined
[3]Xho I restriction site sequence is underlined
[4]From 5' to 3': the overlapping sequence for cpVIII is double underlined; the Spe I restriction site sequence is single underlined; the overlapping sequence for C_H1 3' end is double underlined.
[5]From 5' to 3': Spe I restriction site sequence is single underlined; the overlapping sequence with the 5' end of cpIII is double underlined
[6]From 5' to 3': Nhe I restriction site sequence is single underlined; the overlapping sequence with 3' end of cpIII is double underlined.
[7]From 5' to 3': overlapping sequence with the 3' end of cpIII is double underlined; Nhe I restriction sequence begins with the nucleotide residue "G" at position 4 and extends 5 more residues = GCTAGC.
[8]EcoR I restriction site sequence is single underlined.

To verify amplification of the modified cpVIII membrane anchor domain, the PCR purified DNA products were electrophoresed in a 1% agarose gel. The expected size of the cpVIII was approximately 150 base pairs. The area in the agarose containing the modified cpVIII DNA fragment was isolated from the agarose as described above. The sequence of the isolated modified cpVIII DNA fragment is listed as SEQ. ID. NO. 111. The isolated cpVIII DNA fragment was then admixed with a similarly prepared fragment of modi-

fied Fd as described below in Example 2i in order to form a DNA segment encoding the fusion protein Fd-cpVIII.

cpIII Membrane Anchor: M13mp18 was also used as a source for isolating the gene encoding the membrane anchor domain at cpIII, the sequence of which is listed in the Sequence Listing as SEQ. ID. NO. 112. The amino acid residue sequence of membrane anchor domain cpIII is listed in SEQ. ID. NO. 16. M13mp18 replicative form DNA was prepared as described above and used as a template for PCR for amplifying the mature gene for cpIII membrane anchor domain and incorporating the two cloning sites, Spe I and EcoR I.

The primer pair, G-3(F) (SEQ. ID. NO. 107) and G-3(B) (SEQ. ID. NO. 108) listed in Table 7, was used in PCR as performed above to incorporate Spe I and Nhe I restriction sites. The resultant PCR modified cpIII DNA fragment was verified and purified as described above. The sequence of the PCR modified cpIII membrane anchor domain DNA fragment is listed in the Sequence Listing as SEQ. ID. NO. 113. A second PCR amplification using the primer pairs, Lac-F (SEQ. ID. NO. 109) and Lac-B (SEQ. ID. NO. 110) listed in Table 7, was performed on a separate aliquot of M13mp18 replicative form template DNA. The primers used for this amplification were designed to incorporate an overlapping sequence with the nucleotides encoding the membrane anchor region of cpIII, and the adjacent Nhe I site along with a sequence encoding a LacZ promoter region 5' to an EcoR I restriction site. The reaction and purification of the PCR product performed as described above. The sequence of the resultant PCR modified cpIII DNA fragment having Nho I and EcoR I restriction sites is listed in the Sequence Listing as SEQ. ID. NO. 114.

The products of the first and second PCR amplifications were then recombined at the nucleotides corresponding to cpIII membrane anchor overlap and Nhe I restriction site and subjected to a second round of PCR using the G3-F (SEQ. ID. NO. 107) and Lac-B (SEQ. ID. NO. 110) primer pair to form a recombined PCR DNA fragment product consisting of the following: a 5' Spe I restriction site; a cpIII DNA membrane anchor domain beginning at the nucleotide residue sequence which corresponds to the amino acid residue 198 of the entire mature cpIII protein; an endogenous stop site provided by the membrane anchor at amino acid residue number 112; a Nhe I restriction site, a LacZ promoter sequence; and a 3' EcoR I restriction site. The recombined PCR modified cpIII membrane anchor domain DNA fragment was then restriction digested with Spe I and EcoR I to produce a DNA fragment for directional ligation into a pComb phagemid expression vector prepared in Example 1a(iv) and to form a pCombIII phagemid expression vector as described in Example 1b(ii).

h. Isolation of Anti-NPN Coding V_H DNA Segment

To prepare modified Fd fragments for recombination with the PCR modified cpVIII membrane anchor domain fragment to form a Fd-cpVIII DNA fusion product, PCR amplification as described above was performed using Clone 2b, prepared in Example 2f, as a template. The primers, Hc3 (SEQ. ID. NO. 105) and AK 7 (SEQ. ID. NO. 106), the sequences of which are listed in Table 7, were used in PCR to amplify the Fd portion of the Clone 2b and incorporate Xho I and Spe I cloning sites along with a cpVIII overlapping sequence. The amplified PCR modified Fd product was purified, electrophoresed and isolated from 1% agarose gels as described above. The size of the Fd fragment was 680 base pairs.

i. Preparation of a DNA Segment Encoding a Portion of the Fusion Protein Fd-cpVIII

The purified PCR modified Fd DNA fragment containing cpVIII overlapping nucleotide sequences prepared above was then admixed with the PCR modified cpVIII membrane anchor domain fragment to form an admixture. The fragments in the admixture were allowed to recombine at their complementary regions. The admixture containing the recombined PCR fragments was then subjected to a second round of PCR amplification as described above using the end primer pair AK 6 (SEQ. ID. NO. 104) and Hc3 (SEQ. ID. NO. 105) (Table 7). The corresponding product of the PCR amplification was purified and electrophoresed on agarose gels as described above. The PCR product was determined to be approximately 830 base pairs (Fd=680+150) confirming the fusion of Fd with cpVIII. The sequence of the PCR product linking the Fd sequence with the cpVIII sequence in frame in a 5' to 3' direction is listed as SEQ. ID. NO. 115. The Fd-cpVIII fusion product was then used in directional ligations described in Example 2j for the construction of a pCBAK8-2b dicistronic phagemid expression vector.

j. Construction of pCBAK8-2b Dicistronic Expression Vector

To construct a phagemid vector for the coordinate expression of a Fd-cpVIII fusion protein with kappa light chain, the PCR amplified Fd-cpVIII fusion product prepared in above in Example 2i was first ligated into Clone 2b phagemid expression vector isolated from the NPN combinatorial library prepared in Example 2f. For the ligation, the Fd-cpVIII PCR fusion product was first restriction digested with Xho I and EcoR I. Clone 2b phagemid vector was similarly digested resulting in the removal of the cloning and decapeptide regions. The digested Fd-cpVIII fragment was admixed and ligated into the digested Clone 2b at the cohesive termini generated by Xho I and EcoR I restriction digestion. The ligation resulted in operatively linking the nucleotide residue sequence encoding the Fd-cpVIII polypeptide fusion protein to a second cassette having the nucleotide residue sequences encoding the ribosome binding site, a pelB leader sequence and the kappa light chain already present in Clone 2b to form a dicistronic DNA molecule in the original Clone 2b phagemid expression vector.

E. coli, strain TG1, was then transformed with the phagemid containing the dicistronic DNA molecule and transformants were selected on ampicillin as the original Clone 2b contained an ampicillin selectable resistance marker gene. For high efficiency electro-transformation of *E. coli*, a 1:100 volume of an overnight culture of TG1 cells was inoculated into one liter of L-broth (1% Bacto tryptone, 0.5% Bacto yeast extract, 0.5% NaCl). The cell suspension was maintained at 37° C. with vigorous shaking to a absorbance at 600 nm of 0.5 to 1.0. The cell suspension in log phase growth was then harvested by first chilling the flask on ice for 15 to 30 minutes followed by centrifugation in a cold rotor at 4000 g for 15 minutes to pellet the bacteria. The resultant supernatant was removed and the bacterial cell pellet was resuspended in a total of one liter of cold water to form a cell suspension. The centrifugation and resuspension procedure was repeated two more times and after the final centrifugation, the cell pellet was resuspended in 20 ml of cold 10% glycerol. The resuspended cell suspension was then centrifuged to form a cell pellet. The resultant cell pellet was resuspended to a final volume of 2 to 3 ml in cold 10% glycerol resulting in a cell concentration of 1 to 3×10^{10} cells/ml. For the electro-transformation procedure, 40 ul of the prepared cell suspension was admixed with 1 to 2 ul of phagemid DNA to form a cell-phagemid DNA admixture.

The resultant admixture was mixed and allowed to sit on ice for one minute. An electroporation apparatus, for example a Gene Pulsar, was set a 25 uF and 2.5 kV. The pulse controller was set to 200 ohms. The cell-DNA admixture was transferred to a cold 0.2 cm electroporation cuvette. The cuvette was then placed in the chilled safety chamber and pulsed once at the above settings. To the pulsed admixture, 1 ml of SOC medium was then admixed and the cells were resuspended with a Pasteur pipette (SOC medium was prepared by admixing 2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose). The cells suspension was then transferred to a 17×100 mm polypropylene tube and maintained at 37° C. for one hour. After the maintenance period, the transformed TG1 cells were then plated on ampicillin LB plates for selection of ampicillin resistant colonies containing the phagemid which provided the selectable marker gene.

Ampicillin resistant colonies were selected and analyzed for the correct insert size and expression of Fab. Briefly, DNA minipreps of selected colonies were prepared for the isolation of phagemid DNA. The isolated phagemid DNA from each miniprep was restriction digested with Xho I and EcoR I and the digests were electrophoresed on a 1% agarose gel. Clone AK16 was selected as an 830 bp fragment was visualized on the gels confirming the insertion of the Fd-cpVIII PCR fusion product into digested Clone 2b.

Clone AK16 phagemid was then restriction digested with Xho I and Xba I and the nucleotide residue sequence of the dicistronic DNA molecule encoding the Fd-cpVIII fusion protein, the ribosome binding site and pelB leader sequence for expression of the light chain, a spacer region and the 2b kappa light chain was isolated by agarose gel electrophoresis. The isolated dicistronic DNA fragment was then ligated into a Xho I and Xba I restriction digested pCBAKO expression vector prepared in Example 1c(ii) to form a dicistronic phagemid expression vector designated pCBAK8-2b.

The resultant pCBAK8-2b expression vector consisted of nucleotide residue sequences encoding the following elements: f1 filamentous phage origin of replication; a chloramphenicol acetyl transferase selectable resistance marker gene; an inducible LacZ promoter upstream from the LacZ gene; a multiple cloning site flanked by T3 and T7 polymerase promoters; and the dicistronic DNA molecule (a first cassette consisting of a ribosome binding site, a pelB leader, and a Fd-cpVIII DNA fusion product operatively linked to a second cassette consisting of a second ribosome binding site, a second pelb leader, and a kappa light chain).

k. Construction of pCBAK3-2b Dicistronic Expression Vector

To construct a phagemid vector for the coordinate expression of a Fd-cpIII fusion protein with kappa light chain, the PCR amplified and recombined cpIII membrane anchor prepared in Example 2g having a 5' Spe I and 3' EcoR I restriction site was first directionally ligated into a pComb phagemid expression vector prepared in Example 1a(iv) to form a pCombIII phagemid vector. See Example 1b(ii) for details of vector construction. The resultant pCombIII phagemid vector was then restriction digested with Sac II and Apa I to form an isolated fragment. The resultant isolated fragment containing the expression control sequences and the cpIII sequence was then directionally ligated into a similarly digested pCBAKO phagemid vector prepared in Example 1c(ii) to form a pCBAK3 phagemid expression vector. This vector lacked Fd and kappa light chain sequences.

A phagemid expression vector, pCBAK3-2b, for the expression of a fusion protein and kappa light chain was then constructed. Briefly, the pCBAK3 phagemid expression vector prepared above was first digested with Xho I and Spe I to form a linearized pCBAK3 phagemid expression vector. PCR amplified and modified Fd fragment, prepared in Example 2h containing Xho I and Spe I sites, was subsequently restriction digested with Xho I and Spe I. The resultant Fd fragment was then directionally ligated via cohesive termini into the Xho I and Spe I restriction digested pCBAK3 phagemid expression vector to form a second phagemid expression vector in which the PCR modified Fd fragment was operatively linked in-frame to nucleotide residue sequences encoding cpIII. *E. coli* strain XL1-Blue (Stratagene) was then transformed with the above phagemid vector containing Fd-cpIII. Transformants containing the Fd-cpIII encoding phagemid were selected on chloramphenicol. Phagemid DNA was isolated from chloramphenicol resistant clones and was restriction digested with Sac I and Xba I to form a linearized phagemid expression vector into which a Sac I and Xba I light chain fragment prepared below was directionally ligated.

Phagemid Clone 2b, isolated from the original combinatorial library as described in Example 2a, was restriction digested with Sac I and Xba I to isolate the nucleotide residue sequence encoding the kappa light chain. The isolated kappa light chain sequence was then directionally ligated into the Sac I and Xba I restriction digested phagemid expression vector prepared above containing Fd-cpIII to form the phagemid expression vector, pCBAK3-2b. The resultant vector contained the nucleotide residue sequence of a dicistronic DNA molecule for the coordinate expression of a Fd-cpIII fusion protein with kappa light chain. The resultant phagemid expression vector consisted of nucleotide residue sequences encoding the following elements: f1 filamentous phage origin of replication; a chloramphenicol acetyl transferase selectable resistance marker gene; an inducible LacZ promoter upstream from the LacZ gene; a multiple cloning site flanked by T3 and T7 polymerase promoters; and the dicistronic molecule (a first cassette consisting of a first ribosome binding site and pelB leader operatively linked to Fd-cpIII operatively linked to a second cassette consisting of a second LacZ, a second ribosome binding site, and a second pelB leader operatively linked to a kappa light chain).

XL1-Blue cells were then transformed with the phagemid expression vector pCBAK3-2b. Transformed colonies containing the chloramphenicol resistant phagemids were selected as described above and analyzed for the correct size insert and expression of Fab as described in Example 2j. Following verification of the insert and expression of Fab in the pCBAK3-2b phagemid vector, XL1-Blue cells were then transformed and induced for the expression of Fab antibodies as described in Examples 3 and 4.

3. Expression of Anti-NPN Heterodimer on Phage Surfaces

For expression of antibody Fab directed against NPN on phage surfaces, XL1-Blue cells were separately transformed with the phagemid vectors, pCBAK8-2b and pCBAK3-2b, prepared in Examples 2j and 2k, respectively. The transformants were selected on LB plates containing 30 ug/ml chloramphenicol. Antibiotic resistant colonies were selected for each phagemid transformation and grown in liquid cultures at 37° C. in super broth (super broth was prepared by admixing the following: 20 g 3 [N-Morpholino] propanesulfonic acid (MOPS); 60 g tryptone; 40 g yeast extract; and

2 liter of water; adjust pH to 7.0 with 10 m NaOH) containing 30 ug/ml chloramphenicol and 12.5 ug/ml tetracycline for the respective antibiotic selection of the phagemid and the F' episome. The antibiotic resistant transformed XL1-Blue cells were diluted to an optical density (OD_{600} nm) of 0.4 in super broth. The inducer, isopropyl thiogalactopyranoside (IPTG), was admixed to the bacterial suspension for a final concentration of 1 mM and the admixture was maintained at 37° C. for 1 hour to induce the expression of the fusion protein and kappa light chain from the LacZ promoter. Helper phage, either R408 or VCS M13 (Stratagene), was then admixed to the induced bacterial suspension at a ratio of 10–20 helper phage to 1 transformed bacterial cell to initiate the generation of copies of the sense strand of the phagemid DNA. The admixture containing the helper phage was then maintained for an additional two hours at 37° C. to allow for filamentous bacteriophage assembly wherein the expressed anti-NPN Fab antibodies fused to either bacteriophage membrane anchor domains of cpVIII or cpIII were incorporated into surface of the bacteriophage particles. The bacterial suspension was then centrifuged resulting in a bacterial cell pellet and a supernatant containing phage. The supernatant was removed, collected and assayed as described below for the presence of functional anti-NPN Fab molecules anchored to the phage particles by either cpVIII or cpIII.

4. Assays for Verifying the Presence and Function of Anti-NPN Heterodimer on the Surface of Filamentous Phage

a. Electron Microscopy

To localize functional Fab molecules, the binding to antigen labelled with colloidal gold was studied. Phage containing supernatants and bacterial cells prepared in Example 3 were spotted on formvar Polysciences, Inc., Warrington, Pa.) coated grids affixed onto a solid phase. In some experiments grids were coated with cells and infected with phage in situ. Subsequently grids were blocked with bovine serum albumin (BSA) 1% in PBS at pH 7.2, washed and incubated with 2–7 nanometer (nm) colloidal gold particles coated with BSA-NPN hapten conjugate for a time period sufficient to form a labeled immunoreaction complex. The grids were washed to remove excess gold particles and negatively stained in uranylacetate and visualized by electron microscopy.

Examination of filamentous phage and permeabilized cells producing phage revealed specific labelling of phage or exposed bacterial membranes. Phage were observed to contain 1 to 24 copies of antigen binding sites per particle. Neither helper phage alone nor intact *E. coli* labelled with antigen. Background nonspecific binding was very low. Filamentous phage particles emerging from the *E. coli* surfaces were labelled with antigen as shown in FIG. 9.

The generation of a related phage surface expression vector utilizing cpIII as a fusion partner with Clone 2b, pCBAK3-2b, revealed specific antigen labelling to the phage head but not the column. Additionally human anti-tetanus Fab expressed as a cpIII fusion did not bind to BSA-NPN antigen.

b. Phase Elisa

Microtitration plates were coated with NPN-BSA conjugate (0.1 ml, 1 μg/ml in 0.1M Tris-HCl pH 9.2), and blocked with 1% BSA in PBS. Serial two fold dilutions of pCBAK8-2b derived phage (0.1 ml), prepared in Example 3, were added to the pre-coated microtitration plate and incubated for 3 hours at ambient temperature or 16 hours at 4° C. The plates were washed with PBS and goat anti-kappa alkaline phosphatase conjugate (Fisher Biotech, Pittsburgh, Pa.) added (0.1 ml diluted 1/1000 in PBS containing 0.1% BSA) and incubated for 2 hours at room temperature. The plates were washed in PBS and substrate added (0.1 ml, 1 mg/ml p-nitrophenylphosphate in 0.1M Tris-HCl, pH 9.5, containing 50 mM $MgCl_2$). After incubation at 37° C. for signal development, the optical densities at 400 nm were determined. Competition assays were performed with the addition of increasing amounts of free NPN hapten ranging from zero up to 5 mg/well.

The ELISA assays confirmed the presence of functional antibody Fab. In a two site ELISA on NPN antigen coated plates when probed with anti-mouse kappa chain enzyme conjugate, phage supernatant generated from helper phage infection of cells carrying the pCBAK8-2b construct exhibited expected titration curves with serial two fold dilutions of phage containing antibody. The results of the two-site ELISA are shown in FIG. 10. For a signal to be generated in this assay, the phage particle must (i) have functionally associated Fd and kappa chains and (ii) be multivalent. Specificity of the particle was assessed by inhibiting binding to the plate in the presence of increasing concentrations free hapten. The generated phage particles exhibited binding to solid phase of the ELISA and could be inhibited by addition of hapten as shown in FIG. 11. Complete inhibition was achieved when 5 ng of free NPN hapten was used in the assay. Helper phage did not give a signal in the ELISA.

c. Antigen Specific Precipitation of Phage

Phage supernatant from XL1-Blue was transformed with the pCBAK8-2b dicistronic expression vector prepared in Example 3 (1 ml) was incubated with BSA-NPN conjugate (10 μL 2 mg/ml) for 18 hours at 4° C. The mixture was then pelleted by centrifugation at 3000 rpm on a bench top centrifuge and the appearance of precipitate noted. Helper phage was used as a control. The pellet was washed repeatedly in cold PBS (5×3 ml/wash) and then resuspended in LB (0.5 ml). The solubilized precipitates were added to fresh XLI-Blue cells (0.5 ml of overnight culture), incubated for 1 hour at 37° C. and aliquots plated out on LB agar containing chloramphenicol (30 μg/ml). Colonies were selected randomly. Colony lifts on nitrocellulose were treated with lysozyme to digest the cell wall, briefly treated with chloroform to breakdown the outer membrane, blocked in BSA 1% in PBS and incubated with ^{125}I labelled BSA-NPN antigen. After several washes in PBS (containing 0.05% Tween-20), film was exposed to the washed and dried filter overnight at −70° C. and the autoradiographs were then developed.

Precipitates were obtained with antibody containing phage but not helper phage in the presence of BSA-NPN. In addition, the particles retained infectivity on subsequent incubation with bacterial cells carrying the F' episome and generated 4×10^5 colonies from a single solubilized precipitate.

Additionally, DNA restriction analysis was carried out to determine the presence of heavy and light chain inserts. DNA restriction analysis of the clones revealed the presence of a. Xho and Xba I fragment of 1.4 kb as expected for Fd-cpVIII fusion construct and kappa chain insert.

These results give additional evidence for antigen specificity and multivalency. In addition to providing immunological parameters, this precipitation offers possibilities for facile enrichment of antigen specific phage particles. In principle, phage containing specific antibodies can be highly enriched by precipitation with antigens (which may be cell surface markers, viral, bacterial as well as synthetic molecules). The washed antigen-antibody precipitates can be solubilized by the addition of excess antigen and viable

phage recovered. For the recovery of rare species an immobilized antigen may be used which opens the way for differential affinity elution.

In order to demonstrate the utility of immobilized antigen for the enrichment of clones of defined binding specificity, a panning experiment was performed. An ampicillin resistant phagemid expressing an anti-tetanus Fab as a cpVIII fusion was constructed. Rescue of this clone with helper phage produced phage encoding the ampicillin resistant phagemid which displayed the anti-tetanus Fab on their coat. These phage encoding tetanus specificity were admixed with NPN hapten encoding phage (1:100) and allowed to bind to a microtitration plate coated with tetanus toxoid. Following a one hour maintenance period, the plate was washed extensively and phage were then eluted with a low pH buffer. Infection of XL1-Blue cells in log phase growth and subsequent plating of aliquots on ampicillin and chloramphenicol allowed for direct quantitation of enrichment. Examination of over 1,000 colonies showed that ampicillin resistant colonies derived from the eluted phage exceeded chloramphenicol resistant colonies by 27 to 1. Therefore, panning enriched the phage displaying the anti-tetanus Fab by 2700 fold. This result suggests that a clone of defined specificity present at one part per million will dominate over nonspecific clones following two rounds of panning.

5. Advantages of Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces

A powerful technique for generating and selecting combinatorial Fabs, with 10^{8-9} members, is presented. In the vector described herein, the restriction cloning sites for inserting PCR generated antibody fragments have been retained as previously reported for the lambda vector. The rescue of the genes encoding the antibody Fd and kappa chains is mediated through the utilization of the f1 origin of replication leading to the synthesis and packaging of the positive strand of the vector on co-infection with helper phage. Since the 'mature' virus particle assembles by incorporating the major coat protein around the single stranded DNA as it passes through the inner membrane into the periplasmic space, not only does it capture the genetic information carried on the phagemid vector but also incorporates several copies of functional Fab along the length of the particle. On subsequent infection of hosts cells carrying the F' episome the phagemid confers resistance allowing selection of colonies on the appropriate antibiotic. In essence, the antigen recognition unit has been linked to instructions for its production.

The full power of the earlier combinatorial system could not be fully utilized since screening allowed ready access to only about 0.1–1% of the members. In the phagemid/M13 system similar size libraries are generated and all the members are accessed via affinity selection. Furthermore, unlike the lambda vector which generated monovalent Fabs, this system generates multivalent particles, thus allowing the capture of a wider range of affinities.

The unique phagemid restriction sites permit the recombination of Fd and kappa chains allowing chain replacement or shuffling. The rescue of filamentous single stranded DNA allows rapid sequencing and analysis of the genetic make up of the clone of interest. Indeed it can be envisaged that phage encoding antibody specificity may be enriched by antigen selection prior to DNA sequencing or mutagenesis. The option to further develop an iterative process of mutation followed by selection may allow the rapid generation of high affinity antibodies from germ line sequences. The process may be automated. Setting aside the potential of the system to mimic nature, the phagemid/M13 system would allow a more complete dissection of the antibody response in humans which may yield useful therapeutic and diagnostic reagents.

The membrane anchoring of the heavy chain and the compartmentalization of the kappa chain in the periplasm is the key to expressing this functional dimeric protein. The potential of this system is by no means limited to antibodies and may be extended to any protein recognition system or combination of systems containing multiple members. For example coupling of ligand and effector systems in a high avidity matrix is now possible. In a similar vein a library of ligands can be sorted against a library of receptors.

6. Randomized Mutagenesis of the CDR3 Region of a Heavy Chain Encoding Tetanus Toxoid

a. PCR Mutagenesis with Degenerate Oligonucleotides

To obtain a mutagenized heterodimer of this invention of altered specificity that would no longer recognize TT but would recognize and specifically bind to a new antigen, a method was developed to randomize only the CDR3 region of a heavy chain fragment encoded by a known nucleotide sequence. This approach is schematically diagrammed in FIG. 12 where a representative heavy chain fragment within a phagemid clone, consisting of alternating framework regions (1 through 4) shown by white blocks and complementarity determining regions (CDR) (1 through 3) shown by cross-hatched blocks and the first constant region (CH1), is subjected to two separate rounds of PCR. In the first PCR amplification reaction, the 5' end of the heavy chain beginning at framework 1 and extending to the 3' end of framework 3 is amplified. In the second PCR amplification reaction, the CDR3 region is randomly mutagenized shown by the black box. This is accomplished through the use of a pool of oligonucleotide primers synthesized with a degenerate region sandwiched between and contiguous with conserved framework 3 and 4 region sequences. The resulting amplification products, each having a randomized CDR3 region, begin at the 3' end of framework 3 and extend to the 3' end of the CH1 region. The pool of degenerate oligonucleotide primers have been designed to result in the amplification of products having a 5' end that is complementary to and will overlap with the 3' end of the products of the first PCR reaction product. Thus, the two separate PCR reaction products are pooled and subjected to a third PCR reaction in which the overlapping region between the two products is extended to result in heavy chain having a randomized CDR3 region.

A heavy chain DNA template for use in this invention was available in a clone (a phagemid vector containing heavy and light chain fragments) from a human combinatorial anti-tetanus toxoid (TT) Fab library. This library was constructed in the pCBAK-3 dicistronic expression vector for the expression of a heavy chain-cpIII fusion protein (Fd-cpIII) and a soluble light chain as described for anti-NPN in Example 2k and by Persson et al., *Proc. Natl. Acad. Sci., USA*, 88:2432–2436 (1992) and Barbas et al., *Proc. Natl. Acad. Sci., USA*, 88:7978–7982 (1992). A clone, hereinafter referred to as pCE-TT7E, was expressed as described for anti-NPN heterodimers on phage surfaces in Example 3 and subsequently screened by panning on TT-coated plates as described for anti-NPN in Example 4c . Clone pCE-TT7E exhibited a K_d towards TT on the order of 10^{-7}M and was enriched over nonspecific phage by 10^3-fold as described by Barbas et al., supra. Clone pCE-TT7E, having both heavy and light chain sequences, was used as the template DNA for the randomized mutagenesis of the CDR3 region of the

heavy chain to alter antigen binding specificity as described herein. The sequence of the heavy chain was determined as described in Example 1a(ii). Two separate PCR reactions were performed as illustrated in FIG. 12.

The first PCR reaction resulted in the amplification of the region of the heavy chain fragment in the pC3-TT7E clone beginning of framework region 1 and extending to the end of framework region 3 which is located 5' to CDR3 which is approximately 400 base pairs in length. To amplify this region, the following primer pairs were used. The 5' antisense oligonucleotide primer, FT3X, having the nucleotide sequence 5'-G-CAA-TAA-ACC-CTC-ACT-AAA-GGG-3' (SEQ ID NO 118), hybridized to the non-coding strand of the heavy chain corresponding to the region 5' of and including the beginning of framework 1. The 3' sense oligonucleotide primer, B7EFR3, having the nucleotide sequence 5'-TCT-CGC-ACA-ATA-ATA-CAC-GGC-3' (SEQ ID NO 119), hybridized to the coding strand of the heavy chain corresponding to the 3' end of the framework 3 region. The oligonucleotide primers were synthesized by Research Genetics (Hunstville, Ala.). The PCR reaction was performed in a 100 ul reaction containing one ug of each of oligonucleotide primers FTX3 and B7EFR3, 8 ul 2.5 mM dNTP's (dATP, dCTP, dGTP, dTTP), 1 ul Taq polymerase, 10 ng of template pCE-TT7E, and 10 ul of 10×PCR buffer purchased commercially (Promega Biotech). Two drops of mineral oil were placed on top of the admixture and 35 rounds of PCR amplification in a thermocycler were performed. The amplification cycle consisted of denaturing at 94 C. for one minute, annealing at 50° C. for one minute, followed by extension at 72° C. for two minutes. The resultant PCR amplification products were then gel purified as described in Example 1d and used in an overlap extension PCR reaction with the products of the second PCR reaction, both as described below, to recombine the two products into reconstructed heavy chains containing mutagenized CDR3 regions as illustrated in FIG. 12.

The second PCR reaction resulted in the amplification of the heavy chain from the 3' end of framework region 3 extending to the end of CH1 region which is approximately 390 base pairs in length. To amplify this region, the following primer pairs were used. The 5' anti-sense oligonucleotide primer pool, designated 7ECDR3, had the nucleotide sequence represented by the formula,

5'-GTG-TAT-TAT-TGT-GCG-AGA-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-TGG-GGC-CAA-GGG-A CC-ACG-3'

where N can be A, C, G, or T and S is either C or G (SEQ ID NO 120), wherein the 5' end of the primer pool is complementary to the 3' end of framework 3 represented by the complementary nucleotide sequence of the oligonucleotide primer B73FR3 and the 3' end of the primer pool is complementary to the 5' end of framework 4. The region between the two specified ends of the primer pool is represented by a 48-mer NNS degeneracy which ultimately encodes a diverse population of mutagenized CDR3 regions of 16 amino acid residues in length. The 3' sense oligonucleotide primer, CG1Z, as described by Persson et al., supra, having the nucleotide sequence 5'-GCATGTACTAGTTTTGTCACAAGATTTGGG-3' (SEQ ID NO 121), hybridized to the coding strand of the heavy chain corresponding to the 3' end of the CH1. The second PCR reaction was performed on the pC3-TT7E in a 100 ul reaction as described above containing one ug of each of oligonucleotide primers 7ECDR3 and CG1Z. The resultant PCR amplification product was then gel purified as described above.

One hundred nanograms of gel purified products from the first and second PCR reactions were then admixed with 1 ug each of FTX3 and CG1Z oligonucleotide primers as a primer pair in a final PCR reaction to form a complete heavy chain fragment by overlap extension as illustrated in FIG. 12. The PCR reaction admixture also contained 10 ul 10×PCR buffer, 1 ul Taq polymerase and 8 ul 2.5 mM d NTP's as described above. The PCR reaction was performed as described above. To obtain sufficient quantities of amplification product, 15 identical PCR reactions were performed. The resulting heavy chain fragments beginning at framework 1 and extending to the end of CH1 and having randomly mutagenized CDR3 regions were approximately 790 base pairs in length. The heavy chain fragment amplification products from the 15 reactions were first pooled and then gel purified as described above prior to their incorporation into a phagemid library.

b. Phagemid Library Construction

The resultant gel purified heavy chain fragments prepared in Example 6a were then digested with the restriction enzymes, Xho I and Spe I, as described in Example 2d. The resultant digested heavy chain fragments were subsequently gel purified prior to insertion into the pC3-TT7E phagemid vector clone which was previously digested with the same restriction enzymes to remove the non-mutagenized heavy chain fragment and form a linear vector. Ligation of 640 ng of the heavy chain Xho I/Spe I fragments having mutagenized CDR3 regions into two ug of the linearized pC3-TT7E phagemid vector to form circularized vectors having mutagenized CDR3 regions was performed overnight at room temperature using 10 units of BRL ligase (Gaithersburg, Md.) in BRL ligase buffer in a reaction volume of 150 ul. Five separate ligation reactions were performed to increase the size of the phage library having mutagenized CDR3 regions. Following the ligation reactions, the circularized DNA was precipitated at −20° C. for two hours by the admixture of 2 ul of 20 mg/ml glycogen, 15 ul of 3M sodium acetate at pH 5.2 and 300 ul of ethanol. DNA was then pelleted by microcentrilugation at 4° C. for 15 minutes. The DNA pellet was washed with cold 70% ethanol and dried under vacuum. The pellet was resuspended in 10 ul of water and transformed by electroporation into 300 ul of *E. coli* XL1-Blue cells as described in Example 2k to form a phage library. The total yield from the mutagenesis and transformation procedure described herein was approximately 5×10^7 transformants.

After transformation, to isolate phage on which heterodimer expression has been induced for subsequent panning on target antigens such as fluorescein, 3 ml of SOC medium (SOC was prepared by admixture of 20 g bactotryptone, 5 g yeast extract and 0.5 g NaCl in one liter of water, adjusting the pH to 7.5 and admixing 20 ml of glucose just before use to induce the expression of the Fd-cpIII and light chain heterodimer) was admixed and the culture was shaken at 220 rpm for one hour at 37 C., after which 10 ml of SB (SB was prepared by admixing 30 g tryptone, 20 g yeast extract, and 10 g Mops buffer per liter with pH adjusted to 7) containing 20 ug/ml carbenicillin and 10 ug/ml tetracycline and the admixture was shaken at 300 rpm for an additional hour. This resultant admixture was admixed to 100 ml SB containing 50 ug/ml carbenicillin and 10 ug/ml tetracycline and shaken for one hour, after which helper phage VCSM13 (10^{12} pfu) were admixed and the admixture was shaken for an additional two hours. After this time, 70 ug/ml kanamycin was admixed and maintained at 30° C. overnight. The lower temperature resulted in better heterodimer incorporation on the surface of the phage. The

supernatant was cleared by centrifugation (4000 rpm for 15 minutes in a JA10 rotor at 4° C.). Phage were precipitated by admixture of 4% (w/v) polyethylene glycol 8000 and 3% (w/v) NaCl and maintained on ice for 30 minutes, followed by centrifugation (9000 rpm for 20 minutes in a JA10 rotor at 4° C.). Phage pellets were resuspended in 2 ml of PBS and microcentrifuged for three minutes to pellet debris, transferred to fresh tubes and stored at −20° C. for subsequent screening as described below.

For determining the titering colony forming units (cfu), phage (packaged phagemid) were diluted in SB and 1 ul was used to infect 50 ul of fresh (AOD600=1) *E. coli* XLI-Blue cells grown in SB containing 10 ug/ml tetracycline. Phage and cells were maintained at room temperature for 15 minutes and then directly plated on LB/carbenicillin plates.

c. Selection of Anti-Fluorescein Heterodimers on Phage Surfaces

1) Multiple Pannings of the Phage Library Having Mutagenized CDR3 Regions

The phage library produced in Example 6b having heavy chain fragments with mutagenized CDR3 regions was panned as described herein on a microtiter plate coated with a 50 ug/ml fluorescein-BSA conjugate to screen for anti-fluorescein heterodimers. Fluorescein was conjugated to BSA according to the methods described in "Antibodies: A Laboratory Manual", eds Harlow et al., Cold Spring Harbor Laboratory, 1988.

The panning procedure described was a modification of that originally described by Parmley and Smith (Parmley et al., *Gene*, 73:30–5–318). Two to four wells of a microtiter plate (Costar 3690) were coated overnight at 4° C. with 25 µl of 50 µg/ml antigen prepared above in 0.1M bicarbonate, pH 8.6. The wells were washed twice with water and blocked by completely filling the well with 3% (w/v) bovine serum albumin (BSA) in PBS and incubating the plate at 37° C. for 1 hour. Blocking solution was shaken out. 50 µl of the phage library prepared above (typically 10^{11} cfu) was added to each well, and the plate was incubated for 2 hours at 37° C.

Phage were removed and the plate was washed once with water. Each well was then washed 10 times with TBS/Tween (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% Tween 20) over a period of 1 hour at room temperature —pipetted up and down to wash the well, each time allowing the well to remain completely filled with TBS/Tween between washings. The plate was washed once more with distilled water and adherent phage were eluted by the addition of 50 µl of elution buffer (0.1M HCl adjusted to pH 2.2 with solid glycine, containing 1 mg/ml BSA) to each well and incubation at room temperature for 10 minutes. The elution buffer was pipetted up and down several times, removed, and neutralized with 3 µl of 2M Tris base per 50 µl of elution buffer used. Eluted phage were used to infect 2 ml of fresh (OD_{600}=1) *E. coli* XLI-Blue cells for 15 minutes at room temperature, after which 10 ml of SB containing 20 µg/ml carbenicillin and 10 µg/ml tetracycline was admixed. [Aliquot (20, 10, and 1/10 µl) were removed for plating to determine the number of phage (packaged phagemids) that were eluted from the plate.] The culture was shaken for 1 hour at 37° C., after which it was added to 100 ml of SB containing 50 µg/ml carbenicillin and 10 µg/ml tetracycline and shaken for 1 hour. Then helper phage VCSM13 (10^{12} pfu) were added and the culture was shaken for an additional 2 hours. After this time, 70 µg/ml kanamycin was added and the culture was incubated at 37° C. overnight. Phage preparation and further panning were repeated as described above.

Following each round of panning, the percentage yield of phage must be determined, where % yield—(number of phage eluted/number of phage applied) ×100.

As an alternative to elution with acid, phage bound to the wells of the microtiter plate were eluted by admixing 50 ul of a solution of 10^{-5}M fluorescein diluted in PBS followed by a maintenance period of one hour at 37 C. The solution was then pipetted up and down to wash the wells. The resultant eluate was transferred to 2 ml of fresh *E. coli* XLI-Blue cells for infection as described above for preparing phage and further panning. In subsequent rounds of panning, phage were eluted with 10^{-6}M fluorescein.

The results of the amount of phage that were specifically bound to fluorescein-coated wells over four consecutive rounds of panning and eluted with acid or with fluorescein alone are shown below in Table 8. Comparable yields of phage on which heterodimers were expressed that bound specifically to fluorescein were achieved with either elution protocol. Approximately 20 clones of the 5×10^7 clones resulting from the mutagenesis and transformation exhibited specificity of binding towards fluorescein-coated wells. These data confirm that mutagenesis of the CDR3 region as described in this invention resulted in the altering of a heterodimer which initially specifically bound to TT to one that specifically bound fluorescein.

TABLE 8

	Phage Eluted	
	Acid Elution	Fluorescein Elution
round 1	5.6×10^5/well	4.7×10^5/well
round 2	4.6×10^6/well	5.6×10^5/well
round 3	3.75×10^5/well	1.35×10^6/well
round 4	1.31×10^6/well	4.0×10^6/well

2) Preparation of Soluble Heterodimers for Characterizing Binding Specificity to Fluorescein

In order to further characterize the specificity of the mutagenized heterodimers expressed on the surface of phage as described above, soluble heterodimers were prepared and analyzed in ELISA assays on fluorescein-coated plates, by competive ELISA with increasing concentrations of soluble fluorescein-BSA and also by fluorescence quenching assays. The latter assays were performed as described in "Fluorescein Hapten: An Immunological Probe", ed E. W. Voss, CRC Press, Inc. pp 52–54, 984.

To prepare soluble heterodimers, phagemid DNA from positive clones was isolated and digested with SpeI and NheI Digestion with these enzymes produces compatible cohesive ends. The 4.7-kb DNA fragment lacking the gIII portion was gel-purified (0.6% agarose) and self-ligated. Transformation of *E. coli* XLI-Blue afforded the isolation of recombinants lacking the gIII fragment. Clones were examined for removal of the gIII fragment by XhoI/XbaI digestion, which should yield an 1.6-kb fragment. Clones were grown in 100 ml SB containing 50 µg/ml carbenicillin and 20 mM $MgCl_2$ at 37° C. until an OD_{600} of 0.2 was achieved. IPTG (1 mM) was added and the culture grown overnight at 30° C. (growth at 37° C. provides only a light reduction in heterodimer yield). Cells were pelleted by centrifugation at 4000 rpm for 15 minutes in a JA10 rotor at 4° C. Cells were resuspended in 4 ml PBS containing 34 µg/ml phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication on ice (2–4 minutes at 50% duty). Debris was pelleted by centrifugation at 14,000 rpm in a JA20 rotor at

4° C. for 15 minutes. The supernatant was used directly for ELISA analysis and was stored at −20° C. For the study of a large number of clones, 10-ml cultures provided plenty of heterodimer for analysis. In this case, sonications were performed in 2 ml of buffer.

The soluble heterodimers prepared above were assayed by ELISA. For this assay, 1 ug/well of fluorescein-BSA solution was admixed to individual wells of a microtiter plate and maintained at 4° C. overnight to allow the protein solution to adhere to the walls of the well. After the maintenance period, the wells were washed one time with PBS and thereafter maintained with a solution of 3% BSA to block nonspecific sites on the wells. The plates were maintained at 37° C. for one hour after which time the plates were inverted and shaken to remove the BSA solution. Soluble heterodimers prepared above were then admixed to each well and maintained at 37° C. for one hour to form an immunoreaction products. Following the maintenance period, the wells were washed 10 times with PBS to remove unbound soluble antibody and then maintained with a secondary goat anti-human FAB conjugated to alkaline phosphatase diluted in PBS containing 1% BSA. The wells were maintained at 37° C. for one hour after which the wells were was 10 times with PBS followed by development with p-nitrophenyl phosphate.

Immunoreactive heterodimers as determined in the above ELISA were then analyzed by competition ELISA to determine the affinity of the mutagenized heterodimers. The ELISA was performed as described above with increasing concentrations of soluble fluorescein-BSA ranging in concentration from 10^{-9}M up to 10^{-5}M in concentration admixed in the presence of the soluble heterodimers. Maximal inhibition of binding was achieved at a concentration of 10^{-6}M free antigen with a half-maximal inhibition obtained with approximately 10^{-7}M free antigen. Thus, the mutagenized heterodimers of this invention specifically recognize and bind to fluorescein. Additional experiments were performed to confirm that the mutagenized heterodimers no longer recognized the TT to which they nonmutagenized heterodimer originally bound. Fluorescence quenching assays were also performed to confirm the specificity of binding of the mutagenized heterodimers. Soluble heterodimers prepared from phage that were either eluted with acid or with fluorescein alone were equally effective at binding fluorescein by any of the aforementioned approaches. The invention of mutagenesis of the CDR3 region of the heavy chain of a heterodimer described herein thus resulted in the alteration of binding specificity from TT to fluorescein.

Thus, the above example illustrates a method according to the present invention for mutagenizing the complementarity determining region (CDR) of an immunoglobulin gene, and also illustrates oligonucleotides useful therefor.

In one embodiment, therefore, an oligonucleotide is contemplated that is useful as a primer in a polymerase chain reaction (PCR) for inducing mutagenesis in a complementarity determining region (CDR) of an immunoglobulin gene. The oligonucleotide has 3' and 5' termini and comprises (1) a nucleotide sequence at its 3' terminus capable of hybridizing to a first framework region of an immunoglobulin gene, (2) a nucleotide sequence at its 5' terminus capable of hybridizing to a second framework region of an immunoglobulin gene, and (3) a nucleotide sequence between the 3' and 5' termini adapted for introducing mutations during a PCR into the CDR region between the first and second framework regions of the immunoglobulin gene, thereby mutagenizing the CDR region.

Insofar as immunoglobulin genes have three CDR regions on both the heavy chain and the light chain of an immunoglobulin, each separated by a distinctive framework region, it is to be understood that the above example is readily applicable to introducing mutations into a specific CDR by selection of the above 5' and 3' nucleotide sequences as to hybridize to the framework regions flanking the targeted CDR. Thus the above first and second framework sequences can be the conserved sequences flanking CDR1, CDR2 or CDR3 on either the heavy or light chain. Exemplary and preferred is the CDR3 of the human immunoglobulin heavy chain.

The length of the 3' and 5' terminal nucleotide sequences of a subject mutagenizing oligonucleotide can vary in length as is well known, so long as the length provides a stretch of nucleotides complementary to the target framework sequences as to hybridize thereto. In the case of the 3' terminal nucleotide sequence, it must be of sufficient length and complementarity to the target framework region located 3' to the CDR region to be mutagenized as to hybridize and provide a 3' hydroxyl terminus for initiating a primer extension reaction. In the case of the 5' terminal nucleotide sequence, it must be of sufficient length and complementarity to the target framework region located 5' to the CDR region to be mutagenized as to provide a means for hybridizing in a PCR overlap extension reaction as described above to assemble the complete immunoglobulin heavy or light chain.

Framework regions flanking a CDR are well characterized in the immunological arts, and include known nucleotide sequences or consensus sequences as described elsewhere herein. Where a single, preselected immunoglobulin gene is to be mutagenized, the framework-defined sequences flanking a particular CDR are known, or can be readily determined by nucleotide sequencing protocols. Where a repertoire of immunoglobulin genes are to be mutagenized, the framework-derived sequences are preferably conserved, as described elsewhere herein.

Preferably, the length of the 3' and 5' terminal nucleotide sequences are each at least 6 nucleotides in length, and can be up to 50 or more nucleotides in length, although these lengths are unnecessary to assure accurate and reproducible hybridization. Preferred are lengths in the range of 12 to 30 nucleotides, and typically are about 18 nucleotides.

A particularly preferred framework-defined nucleotide sequence for use as a 3' terminus nucleotide sequence has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122).

A particularly preferred framework-defined nucleotide sequence for use as a 5' terminus nucleotide sequence has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123).

The nucleotide sequence located between the 3' and 5' termini adapted for mutagenizing a CDR can be any nucleotide sequence, insofar as the novel sequence will be incorporated by the above methods. However, the present approach provides a means to produce a large population of mutagenized CDR's in a single PCR reaction by the use of a population of redundant sequences defining randomized or nearly randomized nucleotides in the CDR region to be mutagenized.

A preferred oligonucleotide comprises a nucleotide sequence between the above described 3' and 5' termini that is represented by the formula: $[NNS]_n$ or $[NNK]_n$, wherein N can independently be any nucleotide, where S is G or C, K is G or T, and where n is from 3 to about 24. In preferred embodiments the preferred oligonucleotides have the formula:

5'-GTGTATTATTGTGCGAGA[NNS]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO: 124) and 5'-GTGTATTATTGTGC-GAGA[NNK]$_n$TGGGGCCAAGGGACCACG-3' (SEQ ID NO: 125).

Exemplary and particularly preferred is the oligonucleotide where with the formula $[NNS]_n$ and n is 16, such that the oligonucleotide represents a large population of redundant oligonucleotide sequences.

The invention also contemplates a mutagenesis method for altering the immunological specificity of a cloned immunologlobulin gene. The method provides direct mutagenesis in a preselected CDR of an immunoglobulin gene which comprises subjecting a recombinant DNA molecule (rDNA) containing the cloned immunoglobulin gene having a target CDR to PCR conditions suitable for amplifying a preselected region of the CDR. In the method, the rDNA molecule is subjected to PCR conditions that include a PCR primer oligonucleotide as described above constituting the first primer in a PCR primer pair as is well known to produce an amplified PCR product that is derived from the preselected CDR region but that includes the nucleotide sequences of the PCR primer. The second oligonucleotide in the PCR amplifying conditions can be any PCR primer derived from the immunoglobulin gene to be mutagenized, as described herein.

Preferred are methods using an oligonucleotide of this invention as described above.

The foregoing is intended as illustrative of the present invention but not limiting. Numerous variations and modifications can be effected without departing from the true spirit and scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 125

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 173 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCCGCAAAT TCTATTTCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC TACGGCAGCC 60

GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCC AGGTGAAACT GCTCGAGATT 120

TCTAGACTAG TTACCCGTAC GACGTTCCGG ACTACGGTTC TTAATAGAAT TCG 173

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 173 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCGACGAATT CTATTAAGAA CCGTAGTCCG GAACGTCGTA CGGGTAACTA GTCTAGAAAT 60

CTCGAGCAGT TTCACCTGGG CCATGGCTGG TTGGGCAGCG AGTAATAACA ATCCAGCGGC 120

TGCCGTAGGC AATAGGTATT TCATTATGAC TGTCTCCTTG AAATAGAATT TGC 173

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGAATTCTAA ACTAGTCGCC AAGGAGACAG TCATAATGAA ATACCTATTG CCTACGGCAG 60

CCGCTGGATT GTTATTACTC GCTGCCCAAC CAGCCATGGC CGAGCTCGTC AGTTCTAGAG 120

TTAAGCGGCC G 131

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TCGACGGCCG CTTAACTCTA GAACTGACGA GCTCGGCCAT GGCTGGTTGG GCAGCGAGTA 60

ATAACAATCC AGCGGCTGCC GTAGGCAATA GGTATTTCAT TATGACTGTC TCCTTGGCGA 120

CTAGTTTAGA ATTCAAGCT 139

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met
20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Erwinia carotovora

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

-continued

```
Ala Gln Pro Ala Gln Pro Ala Met Ala
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Erwinia carotovora

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Lys Ser Leu Ile Thr Pro Ile Ala Ala Gly Leu Leu Leu Ala Phe
1               5                  10                  15

Ser Gln Tyr Ser Leu Ala
            20
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Lys Ile Lys Thr Gly Ala Arg Ile Leu Ala Leu Ser Ala Leu Thr
1               5                  10                  15

Thr Met Met Phe Ser Ala Ser Ala Leu Ala Lys Ile
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Met Lys Arg Asn Ile Leu Ala Val Ile Val Pro Ala Leu Leu Val
1               5                  10                  15

Ala Gly Thr Ala Asn Ala Ala Glu
            20
```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

-continued

Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr
1 5 10 15

Pro Val Thr Lys Ala Arg Thr
20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1 5 10 15

Phe Cys Leu Pro Val Phe Ala His Pro
20 25

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Met Ile Thr Leu Arg Lys Leu Pro Leu Ala Val Ala Val Ala Ala
1 5 10 15

Gly Val Met Ser Ala Gln Ala Met Ala Val Asp
20 25

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Lys Ala Thr Lys Leu Val Leu Gly Ala Val Ile Leu Gly Ser Thr
1 5 10 15

Leu Leu Ala Gly Cys Ser
20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

-continued

```
Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
1               5                  10                  15

Val Pro Met Leu Ser Phe Ala
            20
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
1               5                  10                  15

His Ser
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 211 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: protein

(i i i) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Pro Phe Val Cys Glu Tyr Gln Gly Gln Gly Gln Ser Ser Asp Leu Pro
1               5                  10                  15

Gln Pro Pro Val Asn Ala Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly
            20                  25                  30

Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
        35                  40                  45

Ser Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp
    50                  55                  60

Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala
65                  70                  75                  80

Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val
                85                  90                  95

Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser
            100                 105                 110

Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn
        115                 120                 125

Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn
    130                 135                 140

Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg
145                 150                 155                 160

Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys
                165                 170                 175

Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val
            180                 185                 190

Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn
        195                 200                 205
```

-continued

Lys Glu Ser
210

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ala Glu Gly Asp Asp Pro Ala Lys Ala Ala Phe Asn Ser Leu Gln Ala
1 5 10 15

Ser Ala Thr Glu Tyr Ile Gly Tyr Ala Trp Ala Met Val Val Val Ile
20 25 30

Val Gly Ala Thr Ile Gly Ile Lys Leu Phe Lys Lys Phe Thr Ser Lys
35 40 45

Ala Ser
50

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AAUCUUGGAG GCUUUUUUAU GGUUCGUUCU 30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

UAACUAAGGA UGAAAUGCAU GUCUAAGACA 30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

UCCUAGGAGG UUUGACCUAU GCGAGCUUUU 30

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: RNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AUGUACUAAG GAGGUUGUAU GGAACAACGC 30

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGCCGCAAAT TCTATTTCAA GGAGACAGTC AT 32

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATT 36

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

-continued

GTTATTACTC GCTGCCCAAC CAGCCATGGC CC 32

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CAGTTTCACC TGGGCCATGG CTGGTTGGG 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CAGCGAGTAA TAACAATCCA GCGGCTGCCG TAGGCAATAG 40

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTATTTCATT ATGACTGTCT CCTTGAAATA GAATTTGC 38

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AGGTGAAACT GCTCGAGATT TCTAGACTAG TTACCCGTAC 40

(2) INFORMATION FOR SEQ ID NO:29:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGGAACGTCG TACGGGTAAC TAGTCTAGAA ATCTCGAG 38

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GACGTTCCGG ACTACGGTTC TTAATAGAAT TCG 33

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCGACGAATT CTATTAAGAA CCGTAGTC 28

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TGAATTCTAA ACTAGTCGCC AAGGAGACAG TCAT 34

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATT 36

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTTATTACTC GCTGCCCAAC CAGCCATGGC C 31

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GAGCTCGTCA GTTCTAGAGT TAAGCGGCCG 30

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GTATTTCATT ATGACTGTCT CCTTGGCGAC TAGTTTAGAA TTCAAGCT 48

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

-continued

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CAGCGAGTAA TAACAATCCA GCGGCTGCCG TAGGCAATAG 40

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGACGAGCTC GGCCATGGCT GGTTGGG 27

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TCGACGGCCG CTTAACTCTA GAAC 24

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:40:

AGGTSMARCT KCTCGAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AGGTCCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AGGTCCAGCT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AGGTCCAGCT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:44:

AGGTCCAGCT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:45:

AGGTCCAACT GCTCGAGTCT GG 22

-continued

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AGGTCCAACT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:47:

AGGTCCAACT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AGGTCCAACT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGGTNNANCT NCTCGAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO:50:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GCCCAAGGAT GTGCTCACC 19

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CTATTAGAAT TCAACGGTAA CAGTGGTGCC TTGGCCCCA 39

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CTATTAACTA GTAACGGTAA CAGTGGTGCC TTGCCCCA 38

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CTCAGTATGG TGGTTGTGC 19

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

-continued

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GCTACTAGTT TTGATTTCCA CCTTGG 26

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CAGCCATGGC CGACATCCAG ATG 23

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AATTTTACTA GTCACCTTGG TGCTGCTGGC 30

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TATGCAACTA GTACAACCAC AATCCCTGGG CACAATTTT 39

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

-continued

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AGGCTTACTA GTACAATCCC TGGGCACAAT 30

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:59:

CCAGTTCCGA GCTCGTTGTG ACTCAGGAAT CT 32

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CCAGTTCCGA GCTCGTGTTG ACGCAGCCGC CC 32

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CCAGTTCCGA GCTCGTGCTC ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CCAGTTCCGA GCTCCAGATG ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CCAGATGTGA GCTCGTGATG ACCCAGACTC CA 32

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:64:

CCAGATGTGA GCTCGTCATG ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:65:

CCAGATGTGA GCTCTTGATG ACCCAAACTC AA 32

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CCAGATGTGA GCTCGTGATA ACCCAGGATG AA 32

-continued

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC 32

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CCGCCGTCTA GAACACTCAT TCCTGTTGAA GCT 33

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CCGCCGTCTA GAACATTCTG CAGGAGACAG ACT 33

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CCAGTTCCGA GCTCGTGATG ACACAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

-continued

(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GCGCCGTCTA GAATTAACAC TCATTCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CTATTAACTA GTAACGGTAA CAGTGGTGCC TTGCCCCA 38

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:73:

AGGCTTACTA GTACAATCCC TGGGCACAAT 30

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:74:

GCCGCTCTAG AACACTCATT CCTGTTGAA 29

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:75:

AGGTNNANCT NCTCGAGTCT GC 22

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:76:

AGGTNNANCT NCTCGAGTCA GC 22

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GTGCCAGATG TGAGCTCGTG ATGACCCAGT CTCCA 35

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:78:

TCCTTCTAGA TTACTAACAC TCTCCCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GCATTCTAGA CTATTATGAA CATTCTGTAG GGGC 34

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:80:

CTGCACAGGG TCCTGGGCCG AGCTCGTGGT GACTCAG 37

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:81:

AGNTGCANNT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GTGGGCATGT GTGAGTTGTG TCACTAGTTG GGGTTTTGAG CTC 43

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:83:

-continued

AGCATCACTA GTACAAGATT TGGGCTC 27

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:84:

AGGTGCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:85:

AGGTGCAGCT GCTCGAGTCG GG 22

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:86:

AGGTGCAACT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:87:

AGGTGCAACT GCTCGAGTCG GG 22

-continued

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:88:

TCCTTCTAGA TTACTAACAC TCTCCCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:89:

CTGCACAGGG TCCTGGGCCG AGCTCGTGGT GACTCAG 37

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:90:

GCATTCTAGA CTATTAACAT TCTGTAGGGG C 31

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:91:

ACCCAAGGAC ACCCTCATG 19

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs

-continued

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CTCAGTATGG TGGTTGTGC 19

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GTCTCACTAG TCTCCACCAA GGGCCCATCG GTC 33

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:94:

ATATACTAGT GAGACAGTGA CCAGGGTTCC TTGGCCCCA 39

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:95:

ACGTCTAGAT TCCACCTTGG TCCC 24

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GCATACTAGT CTATTAACAT TCTGTAGGGG C 31

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CCGGAATTCT TATCATTTAC CCGGAGA 27

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:98:

TCTGCACTAG TTGGAATGGG CACATGCAG 29

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 798 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GGCCGCAAAT TCTATTTCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC TACGGCAGCC 60

GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCC AGGTGAAACT GCTCGAGTCA 120

GGACCTGGCC TCGTGAAACC TTCTCAGTCT CTGTCTCTCA CCTGCTCTGT CACTGACTAC 180

TCCATCACCA GTGCTTATTA CTGGAACTGG ATCCGGCAGT TTCCAGGAAA CAAACTGGAA 240

TGGATGGGCT ACATAAGCTA CGACGGTGTC AATAAGTATG ATCCATCTCT CAAGAATCGA 300

ATCTCCATCA CTCGTGACAC ATCTAACAAT CAGTTTTTCC AGAAGTTGAT TTCTGTGACT 360

TCTGAGGACA CAGGAACATA TGACTGTTCA AGAGGGACTA GGGCCTCTGC TATGGACTAC 420

TGGGGTCAAG GAATTTCAGT CACCGTCTCC TCAGCCAAAA CGACACCCCC ATCTGTCTAT 480

-continued

CCACTGGCCC CTGGATCTGC TGCCCAAACT AACTCCATGG TGACCCTGGG ATGCCTGGTC 540

AAGGGCTATT TCCCTGAGCC AGTGACAGTG ACCTGGAACT CTGGATCCCT GTCCAGCGGT 600

GTGCACACCT TCCCAGCTGT CCTGCAGTCT GACCTCTACA CTCTGAGCAG CTCAGTGACT 660

GTCCCCTCCA GCCCTCGGCC CAGCGAGACC GTCACCTGCA ACGTTGCCCA CCCGGCCAGC 720

AGCACCAAGG TGGACAAGAA AATTGTGCCC AGGGATTGTA CTAGTTACCC GTACGACGTT 780

CCGGACTACG GTTCTTAA 798

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 194 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TGAATTCTAA ACTAGTCGCC AAGGAGACAG TCATAATGAA ATACCTATTG CCTACGGCAG 60

CCGCTGGATT GTTACTCGCT GCCCAACCAG CCATGGCCGA GCTCCAGATG ACCCAGTCTC 120

CAGCCTCCCT ATCTGCATCT GTGGGAGAAA CTGTCACCAT CACATGTCGA TCAAGTGAGA 180

ATATTACAAT TACT 194

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 333 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CTGATGCTGC ACCAACTGTA TCCATCTTCC CACCATCCAG TGAGCAGTTA ACATCTGGAG 60

GTGCCTCAGT CGTGTGCTTC TTGAACAACT TCTACCCCAA AGACTACAAT GTCAAGGGGA 120

AGATTGATGG CAGTGAACGA CAAAATGGCG TCCTGAACAG TTGGACTGAT CAGGACAGCA 180

AAGACAGCAC CTACAGCATG AGCAGCACCC TCACGTTGAC CAAGGACGAG TATGAACGAC 240

ATAACAGCTA TACCTGTGAT GCCACTCACA AGACATCAAC TTCACCCATT GTCAAGAGCT 300

TCAACAGGAA TGAGTGTTAA TTCTAGACGG CGC 333

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:102:

GCTGAGGGTG ACGATCCCGC AAAAGCGGCC TTTAACTCCC TGCAAGCCTC AGCGACCGAA 60

TATATCGGTT ATGCGTGGGC GATGGTTGTT GTCATTGTCG GCGCAACTAT CGGTATCAAG 120

CTGTTTAAGA AATTCACCTC GAAAGCAAGC 150

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GTGCCCAGGG ATTGTACTAG TGCTGAGGGT GACGAT 36

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:104:

ACTCGAATTC TATCAGCTTG CTTTCGAGGT GAA 33

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:105:

AGGTCCAGCT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GTCACCCTCA GCACTAGTAC AATCCCTGGG CAC 33

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GAGACGACTA GTGGTGGCGG TGGCTCTCCA TTCGTTTGTG AATATCAA 48

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:108:

TTACTAGCTA GCATAATAAC GGAATACCCA AAAGAACTGG 40

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:109:

TATGCTAGCT AGTAACACGA CAGGTTTCCC GACTGG 36

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:110:

-continued

ACCGAGCTCG AATTCGTAAT CATGGTC 27

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GTGCCCAGGG ATTGTACTAG TGCTGAGGGT GACGATCCCG CAAAAGCGGC CTTTAACTCC 60

CTGCAAGCCT CAGCGACCGA ATATATCGGT TATGCGTGGG CGATGGTTGT TGTCATTGTC 120

GGCGCAACTA TCGGTATCAA GCTGTTTAAG AAATTCACCT CGAAAGCAAG CTGATAGAAT 180

TCGAGT 186

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 666 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:112:

CCATTCGTTT GTGAATATCA AGGCCAAGGC CAATCGTCTG ACCTGCCTCA ACCTCCTGTC 60

AATGCTGGCG GCGGCTCTGG TGGTGGTTCT GGTGGCGGCT CTGAGGGTGG TGGCTCTGAG 120

GGTGGCGGTT CTGAGGGTGG CGGCTCTGAG GGAGGCGGTT CCGGTGGTGG CTCTGGTTCC 180

GGTGATTTTG ATTATGAAAA GATGGCAAAC GCTAATAAGG GGGCTATGAC CGAAAATGCC 240

GATGAAAACG CGCTACAGTC TGACGCTAAA GGCAAACTTG ATTCTGTCGC TACTGATTAC 300

GGTGCTGCTA TCGATGGTTT CATTGGTGAC GTTTCCGGCC TTGCTAATGG TAATGGTGCT 360

ACTGGTGATT TTGCTGGCTC TAATTCCCAA ATGGCTCAAG TCGGTGACGG TGATAATTCA 420

CCTTTAATGA ATAATTTCCG TCAATATTTA CCTTCCCTCC CTCAATCGGT TGAATGTCGC 480

CCTTTTGTCT TTAGCGCTGG TAAACCATAT GAATTTTCTA TTGATTGTGA CAAAATAAAC 540

TTATTCGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 600

TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 660

TATTAT 666

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 708 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:113:

GAGACGACTA GTGGTGGCGG TGGCTCTCCA TTCGTTTGTG AATATCAAGG CCAAGGCCAA 60

TCGTCTGACC TGCCTCAACC TCCTGTCAAT GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT 120

GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA 180

GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT GATTTTGATT ATGAAAAGAT GGCAAACGCT 240

AATAAGGGGG CTATGACCGA AAATGCCGAT GAAAACGCGC TACAGTCTGA CGCTAAAGGC 300

AAACTTGATT CTGTCGCTAC TGATTACGGT GCTGCTATCG ATGGTTTCAT TGGTGACGTT 360

TCCGGCCTTG CTAATGGTAA TGGTGCTACT GGTGATTTTG CTGGCTCTAA TTCCCAAATG 420

GCTCAAGTCG GTGACGGTGA TAATTCACCT TTAATGAATA ATTTCCGTCA ATATTTACCT 480

TCCCTCCCTC AATCGGTTGA ATGTCGCCCT TTTGTCTTTA GCGCTGGTAA ACCATATGAA 540

TTTTCTATTG ATTGTGACAA AATAAACTTA TTCCGTGGTG TCTTTGCGTT TCTTTTATAT 600

GTTGCCACCT TTATGTATGT ATTTTCTACG TTTGCTAACA TACTGCGTAA TAAGGAGTCT 660

TAATCATGCC AGTTCTTTTG GGTATTCCGT TATTATGCTA GCTAGTAA 708

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 201 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:114:

TATGCTAGCT AGTAACACGA CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA 60

ATTAATGTGA GTTAGCTCAC TCATTAGGCA CCCCAGGCTT TACACTTTAT GCTTCCGGCT 120

CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTCACA CAGGAAACAG CTATGACCAT 180

GATTACGAAT TCGAGCTCGG T 201

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 830 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:115:

AGGTCCAGCT TCTCGAGTCT GGACCTGGCC TCGTGAAACC TTCTCAGTCT CTGTCTCTCA 60

CCTGCTCTGT CACTGACTAC TCCATCACCA GTGCTTATTA CTGGAACTGG ATCCGGCAGT 120

TTCCAGGAAA CAAACTGGAA TGGATGGGCT ACATAAGCTA CGACGGTGTC AATAAGTATG 180

ATCCATCTCT CAAGAATCGA ATCTCCATCA CTCGTGACAC ATCTAACAAT CAGTTTTTCC 240

AGAAGTTGAT TTCTGTGACT TCTGAGGACA CAGGAACATA TGACTGTTCA AGAGGGACTA 300

-continued

GGGCCTCTGC TATGGACTAC TGGGGTCAAG GAATTTCAGT CACCGTCTCC TCAGCCAAAA 360

CGACACCCCC ATCTGTCTAT CCACTGGCCC CTGGATCTGC TGCCCAAACT AACTCCATGG 420

TGACCCTGGG ATGCCTGGTC AAGGGCTATT TCCCTGAGCC AGTGACAGTG ACCTGGAACT 480

CTGGATCCCT GTCCAGCGGT GTGCACACCT TCCCAGCTGT CCTGCAGTCT GACCTCTACA 540

CTCTGAGCAG CTCAGTGACT GTCCCCTCCA GCCCTCGGCC CAGCGAGACC GTCACCTGCA 600

ACGTTGCCCA CCCGGCCAGC AGCACCAAGG TGGACAAGAA AATTGTGCCC AGGGATTGTA 660

CTAGTGCTGA GGGTGACGAT CCCGCAAAAG CGGCCTTTAA CTCCCTGCAA GCCTCAGCGA 720

CCGAATATAT CGGTTATGCG TGGGCGATGG TTGTTGTCAT TGTCGGCGCA ACTATCGGTA 780

TCAAGCTGTT TAAGAAATTC ACCTCGAAAG CAAGCTGATA GAATTCGAGT 830

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 260 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:116:

ATGAAATACC TATTGCCTAC GGCAGCCGCT GGATTGTTAT TACTCGCTGC CCAACCAGCC 60

ATGGCCCAGG TGAAACTGCT CGAGATTTCT AGACTAGTGC TGAGGGTGAC GATCCCGCAA 120

AAGCGGCCTT TAACTCCCTG CAAGCCTCAG CGACCGAATA TATCGGTTAT GCGTGGGCGA 180

TGGTTGTTGT CATTGTCGGC GCAACTATCG GTATCAAGCT GTTTAAGAAA TTCACCTCGA 240

AAGCAAGCTG ATAGAATTCG 260

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:117:

GTACGCGCCC TGTAGCGGCG CATTAAGCGC GGCGGGTGTG GTGGTTACGC GCAGCGTGAC 60

CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTTGCT TTCTTCCCTT CCTTTCTCGC 120

CACGTTCGCC GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG GGTTCCGATT 180

TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGTT CACGTAGTGG 240

GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT TCTTTAATAG 300

TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC TCGGTCTATT CTTTTGATTT 360

ATAAGGGATT TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT AACAAAAATT 420

TAACGCGAAT TTTAACAAAA TATTAACGTT TACAATTTAA A 461

-continued

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:118:

GCAATAAACC CTCACTAAAG GG 22

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:119:

TCTCGCACAA TAATACACGG C 21

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 84 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:120:

GTGTATTATT GTGCGAGANN SNNSNNSNNS NNSNNSNNSN NSNNSNNSNN SNNSNNSNNS 60

NNSNNSTGGG GCCAAGGGAC CACG 84

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:121:

GCATGTACTA GTTTTGTCAC AAGATTTGGG 30

(2) INFORMATION FOR SEQ ID NO:122:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:122:

TGGGGCCAAG GGACCACG 18

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:123:

GTGTATTATT GTGCGAGA 18

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(i x) FEATURE:
(A) NAME/KEY: repeat_region
(B) LOCATION: 19..21
(D) OTHER INFORMATION: /rpt_type="tandem"
/ note="NNS can be repeated from 3 to about 24
times."

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:124:

GTGTATTATT GTGCGAGANN STGGGGCCAA GGGACCACG 39

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(i x) FEATURE:
(A) NAME/KEY: repeat_region
(B) LOCATION: 19..21
(D) OTHER INFORMATION: /rpt_type="tandem"

-continued

/ note="NNK can be repeated from 3 to about 24 times."

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:125:

GTGTATTATT GTGCGAGANN KTGGGGCCAA GGGACCACG 39

What is claimed is:

1. A method for producing a library of complementarity determining region (CDR)-mutagenized phage-displayed immunoglobulin heterodimers, the method comprising the steps of:

1) amplifying a CDR portion of a template immunoglobulin variable domain gene selected from the group consisting of a template immunoglobulin heavy chain variable domain gene and a template immunoglobulin light chain variable domain gene, wherein said template immunoglobulin heavy and light chain genes have a framework region and said CDR portion and encode respective heavy and light chain variable domain polypeptides, and wherein said amplifying is by polymerase chain reaction (PCR) using a PCR primer oligonucleotide for mutagenizing a preselected nucleotide region in said CDR portion, thereby forming a library of amplified CDR-mutagenized immunoglobulin gene fragments, said PCR primer oligonucleotide having 3' and 5' termini and comprising:

a) a nucleotide sequence at the 3' terminus capable of hybridizing to a first framework region of said selected template immunoglobulin variable domain gene;

b) a nucleotide sequence at the 5' terminus capable of hybridizing to a second framework region of said selected template immunoglobulin variable domain gene; and

c) a nucleotide sequence between the 3' and 5' termini according to the formula selected from the group consisting of:

$[NNS]_n$ and $[NNK]_n$,

wherein N is independently any nucleotide, S is G or C and K is G or T, and n is 3 to about 24, the 3' and 5' terminal nucleotide sequences having a length of about 6 to 50 nucleotides, or an oligonucleotide having a sequence complementary thereto;

2) inserting individual members of the library of amplified CDR-mutagenized immunoglobulin gene fragments formed in step (1) into a dicistronic phagemid expression vector comprising immunoglobulin heavy and light chain variable domain genes that lack the immunoglobulin gene portion corresponding to the fragment to be inserted, wherein upon insertion said vector is capable of expressing heavy and light chain variable domain polypeptides encoded by said vector, thereby forming a library of dicistronic expression vectors containing amplified CDR-mutagenized immunoglobulin gene fragments; and

3) expressing said immunoglobulin heavy and light chain genes in the library of dicistronic expression vectors formed in step (2) whereby said encoded heavy and light chain variable domain polypeptides assemble on the surface of a phage to form a phage-displayed immunoglobulin heterodimer, thereby producing a library of CDR-mutagenized phage-displayed immunoglobulin heterodimers.

2. The method of claim 1 wherein said 3' terminus has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122), or an oligonucleotide having a sequence complementary thereto.

3. The method of claim 1 wherein said 5' terminus has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123) or an oligonucleotide having a sequence complementary thereto.

4. The method of claim 1 wherein said template immunoglobulin heavy and light chain genes are obtained from a human.

5. The method of claim 1 wherein said CDR portion is CDR3.

6. The method of claim 1 wherein said formula is:

5'-GTGTATTATTGTGCGAGA$[NNS]_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 124).

7. The method of claim 1 wherein n is 16 in the formula $[NNS]_n$ (SEQ ID NO 120).

8. The method of claim 1 wherein said formula is:

5'-GTGTATTATTGTGCGAGA$[NNK]_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 125).

9. A method for producing a complementarity determining region (CDR)-mutagenized phage-displayed immunoglobulin heterodimer with altered antigen binding specificity, the method comprising the steps of:

1) amplifying a CDR portion of a template immunoglobulin variable domain gene selected from the group consisting of a template immunoglobulin heavy chain variable domain gene and a template immunoglobulin light chain variable domain gene, wherein said template immunoglobulin heavy and light chain genes have a framework region and said CDR portion and encode respective heavy and light chain variable domain polypeptides having a preselected antigen binding specificity to a first antigen, and wherein said amplifying is by polymerase chain reaction (PCR) using a PCR primer oligonucleotide for mutagenizing a preselected nucleotide region in said CDR portion to alter said preselected antigen binding specificity, thereby forming a library of amplified CDR-mutagenized immunoglobulin gene fragments, said PCR primer oligonucleotide having 3' and 5' termini and comprising:

a) a nucleotide sequence at the 3' terminus capable of hybridizing to a first framework region of said selected template immunoglobulin variable domain gene;

b) a nucleotide sequence at the 5' terminus capable of hybridizing to a second framework region of said selected template immunoglobulin variable domain gene; and

c) a nucleotide sequence between the 3' and 5' termini according to the formula selected from the group consisting of:

$[NNS]_n$ and $[NNK]_n$,

wherein N is independently any nucleotide, S is G or C and K is G or T, and n is 3 to about 24, the 3' and 5' terminal nucleotide sequences having a length of about 6 to 50 nucleotides, or an oligonucleotide having a sequence complementary thereto;

2) inserting individual members of the library of amplified CDR-mutagenized immunoglobulin gene fragments formed in step (1) into a dicistronic phagemid expression vector comprising immunoglobulin heavy and light chain variable domain genes that lack the immunoglobulin gene portion corresponding to the fragment to be inserted, wherein upon insertion said vector is capable of expressing heavy and light chain variable domain polypeptides encoded by said vector, thereby forming a library of dicistronic expression vectors containing amplified CDR-mutagenized immunoglobulin gene fragments;

3) expressing said immunoglobulin heavy and light chain genes in the library of dicistronic expression vectors formed in step (2) whereby said encoded heavy and light chain variable domain polypeptides assemble on the surface of a phage to form a phage-displayed immunoglobulin heterodimer, thereby producing a library of CDR-mutagenized phage-displayed immunoglobulin heterodimers; and

4) immunoreacting members of the library of CDR-mutagenized phage-displayed immunoglobulin heterodimers formed in step (3) on a preselected second antigen, said second antigen being different than said first antigen to allow for selection of a CDR-mutagenized phage-displayed immunoglobulin heterodimer with altered antigen binding specificity.

10. The method of claim **9** wherein said 3' terminus has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122), or an oligonucleotide having a sequence complementary thereto.

11. The method of claim **9** wherein said 5' terminus has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123) or an oligonucleotide having a sequence complementary thereto.

12. The method of claim **9** wherein said template immunoglobulin heavy and light chain genes are obtained from a human.

13. The method of claim **9** wherein said CDR portion is CDR3.

14. The method of claim **9** wherein said formula is

5'-GTGTATTATTGTGCGAGA[NNS]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 124).

15. The method of claim **9** wherein n is 16in the formula [NNS]$_n$ (SEQ ID NO 120).

16. The method of claim **9** wherein said formula is:

5'-GTGTATTATTGTGCGAGA[NNK]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 125).

17. A method for producing a soluble complementarity determining region (CDR)-mutagenized immunoglobulin heterodimer with altered immunoreactivity to a preselected antigen, the method comprising the steps of:

1) amplifying a CDR portion of a template immunoglobulin variable domain gene selected from the group consisting of a template immunoglobulin heavy chain variable domain gene and a template immunoglobulin light chain variable domain gene, wherein said template immunoglobulin heavy and light chain genes have a framework region and said CDR portion and encode respective heavy and light chain variable domain polypeptides immunoreactive with a preselected antigen, and wherein said amplifying is by polymerase chain reaction (PCR) using a PCR primer oligonucleotide for mutagenizing a preselected nucleotide region in said CDR portion to alter said immunoreactivity of said immunoglobulin heterodimer to the preselected antigen, thereby forming a library of amplified CDR-mutagenized immunoglobulin gene fragments, said PCR primer oligonucleotide having 3' and 5' termini and comprising:

a) a nucleotide sequence at the 3' terminus capable of hybridizing to a first framework region of said selected template immunoglobulin variable domain gene;

b) a nucleotide sequence at the 5' terminus capable of hybridizing to a second framework region of said selected template immunoglobulin variable domain gene; and

c) a nucleotide sequence between the 3' and 5' termini according to the formula selected from the group consisting of:

[NNS]$_n$ and [NNK]$_n$

wherein N is independently any nucleotide, S is G or C and K is G or T, and n is 3 to about 24, the 3' and 5' terminal nucleotide sequences having a length of about 6 to 50 nucleotides, or an oligonucleotide having a sequence complementary thereto;

2) inserting individual members of the library of amplified CDR-mutagenized immunoglobulin gene fragments formed in step (1) into a dicistronic phagemid expression vector comprising immunoglobulin heavy and light chain variable domain genes that lack the immunoglobulin gene portion corresponding to the fragment to be inserted, wherein upon insertion said vector is capable of expressing heavy and light chain variable domain polypeptides encoded by said vector, thereby forming a library of dicistronic expression vectors containing amplified CDR-mutagenized immunoglobulin gene fragments;

3) expressing said immunoglobulin heavy and light chain genes in the library of dicistronic expression vectors formed in step (2) whereby said encoded heavy and light chain variable domain polypeptides assemble on the surface of a phage to form a phage-displayed immunoglobulin heterodimer, thereby producing a library of CDR-mutagenized phage-displayed immunoglobulin heterodimers;

4) immunoreacting members of the library of CDR-mutagenized phage-displayed immunoglobulin heterodimers produced in step (5) on said preselected antigen to isolate an immunoglobulin heterodimer having altered immunoreactivity;

5) isolating said immunoreacted CDR-mutagenized phage-displayed immunoglobulin heterodimer obtained in step (4);

6) producing a soluble form of said immunoreacted form of immunoreacted CDR-mutagenized phage-displayed immunoglobulin heterodimer isolated in step (5); and

7) assaying said soluble form of immunoreacted CDR-mutagenized immunoglobulin heterodimer prepared in step (6) to identify a CDR-mutagenized immunoglobulin heterodimer with altered immunoreactivity to the preselected antigen.

18. The method of claim **17** wherein said assaying is determined by an increase in affinity to the preselected antigen.

19. The method of claim 18 wherein the affinity is greater than $10^{-5}M^{-1}$ dissociation constant (K_d).

20. The method of claim 17 wherein said 3' terminus has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122), or an oligonucleotide having a sequence complementary thereto.

21. The method of claim 17 wherein said 5' terminus has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123) or an oligonucleotide having a sequence complementary thereto.

22. The method of claim 17 wherein said template immunoglobulin heavy and light chain genes are obtained from a human.

23. The method of claim 17 wherein said CDR portion is CDR3.

24. The method of claim 17 wherein said formula is

5'-GTGTATTATTGTGCGAGA[NNS]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 124).

25. The method of claim 17 wherein n is 16in the formula [NNS]$_n$ (SEQ ID NO 120).

26. The method of claim 17 wherein said formula is:

5'-GTGTATTATTGTGCGAGA[NNK]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 125).

* * * * *

In vitro evolution of a T cell receptor with high affinity for peptide/MHC

Phillip D. Holler*, Philmore O. Holman*, Eric V. Shusta†, Sean O'Herrin‡, K. Dane Wittrup†§, and David M. Kranz*¶

Departments of *Biochemistry and †Chemical Engineering, University of Illinois, Urbana, IL 61801; and ‡Ben May Institute for Cancer Research, University of Chicago, Chicago, IL 60637

Communicated by Herman N. Eisen, Massachusetts Institute of Technology, Cambridge, MA, February 23, 2000 (received for review December 15, 1999)

T cell receptors (TCRs) exhibit genetic and structural diversity similar to antibodies, but they have binding affinities that are several orders of magnitude lower. It has been suggested that TCRs undergo selection *in vivo* to maintain lower affinities. Here, we show that there is not an inherent genetic or structural limitation on higher affinity. Higher-affinity TCR variants were generated in the absence of *in vivo* selective pressures by using yeast display and selection from a library of Vα CDR3 mutants. Selected mutants had greater than 100-fold higher affinity ($K_D \approx 9$ nM) for the peptide/MHC ligand while retaining a high degree of peptide specificity. Among the high-affinity TCR mutants, a strong preference was found for CDR3α that contained Pro or Gly residues. Finally, unlike the wild-type TCR, a soluble monomeric form of a high-affinity TCR was capable of directly detecting peptide/MHC complexes on antigen-presenting cells. These findings prove that affinity maturation of TCRs is possible and suggest a strategy for engineering TCRs that can be used in targeting specific peptide/MHC complexes for diagnostic and therapeutic purposes.

T cells recognize a foreign peptide bound to the MHC product through the αβ heterodimeric receptor. The T cell receptor TCR) repertoire has extensive diversity created by the same ,ene rearrangement mechanisms used in antibody heavy- and light-chain genes (1). Most of the diversity is generated at the junctions of V and J (or diversity, D) regions that encode the complementarity-determining region three (CDR3) of the α and β chains (2). However, TCRs do not undergo somatic point mutations as do antibodies, and perhaps not coincidentally, TCRs also do not undergo the same extent of affinity maturation as antibodies. TCRs appear to have affinities that range from 10^5 to 10^7 M^{-1} whereas antibodies have affinities that range from 10^5 to 10^{10} M^{-1} (3, 4).

Whereas the absence of somatic mutation in TCRs may be associated with lower affinities, it has also been argued that there is not a selective advantage for a TCR to have higher affinity (5–7). In fact, the serial-triggering (6) and kinetic proofreading (7) models of T cell activation both suggest that very slow off-rates (associated with higher affinity) would be detrimental to the signaling process. On the other hand, the fastest off-rates that have been measurable have been associated with altered pMHC that exhibits antagonist activity (8–11). Whereas the narrow range of natural TCR affinities has provided some evidence for the relationships between off-rates and agonist/antagonist activity, there are also examples that appear to be inconsistent with these hypotheses (12, 13).

There are other possible explanations for why the T cell system maintains relatively low TCR:pMHC affinities *in vivo*. Peptides bound within the MHC groove display limited accessible surface (14), which may in turn limit the amount of free energy that can be generated in the interaction. On the other hand, raising the affinity of a TCR by directing the free energy toward the MHC helices would presumably lead to thymic deletion during negative selection (15). Even if such higher-affinity TCR could escape thymic deletion, they would likely not maintain the peptide specificity required for T cell responses.

It has not been possible to directly test these possibilities because the generation of TCRs with affinities above 10^7 M^{-1} has not been accomplished. In addition to allowing a kinetic basis of T cell triggering, high-affinity TCRs could be used to more easily explore the role of peptide in pMHC recognition, and as quantitative probes for the expression of pMHC on various target cells. Because *in vivo* selection schemes have not yielded TCRs with the intrinsic binding affinities of affinity-matured antibodies, in this report, we have used an *in vitro* method for the directed evolution of high-affinity TCRs. The method relies on the expression of a library of mutant single-chain (Vβ-linker-Vα) TCRs on the surface of yeast, as a fusion to the surface protein Aga-2 (16, 17). Our previous studies have shown that the yeast display system could be used to engineer variants of the 2C single-chain TCR (scTCR) that were more thermally stable and secreted at higher levels (17, 18). The stability mutants were isolated by subjecting the entire TCR gene to random mutagenesis and selecting for increased surface levels with anti-TCR antibodies (17). The mutations that increased stability resided at the Vα:Vβ interface or on the outside surface of Vβ in a region not involved in pMHC binding. To isolate TCR with higher affinity for pMHC, in the present study, we mutated only the CDR3α loop, which is at the center of the pMHC-binding site (19). Our efforts were guided by previous findings that this region contributed minimal binding free energy to the interaction of the 2C TCR with the pMHC ligand QL9/L^d (20), suggesting that productive interactions might be improved by focusing on this region. Remarkably, selection from a relatively small library (10^5 mutants) yielded many different TCRs with up to 100-fold increased affinity for QL9/L^d. The high-affinity TCRs retained a high degree of peptide specificity although there was some variation in fine specificity among the mutants. These findings suggest that the *in vitro* evolution process described here can be used to isolate TCRs with specificities that one defines by selection with appropriate pMHC ligands.

The high-affinity receptors in this study were derived by variation at the VJ junction, the same process that operates very effectively *in vivo* through gene rearrangements in T cells (2). The fact that we could readily isolate a diverse set of high-affinity TCR *in vitro* indicates that there is not a genetic or structural limitation to high-affinity receptors. This supports the view that inherently low affinities of TCRs found *in vivo* are caused by a lack of selection for higher affinity and perhaps a selection for lower affinity (5–7). Finally, the high-affinity TCR were used in monomeric form to detect pMHC on the surface of target cells,

Abbreviations: TCR, T cell receptor; scTCR, single-chain T cell receptor; CDR3, complementarity-determining regions three.

§Present address: Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

¶To whom reprint requests should be addressed. E-mail: d-kranz@uiuc.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.080078297. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.080078297

IMMUNOLOGY

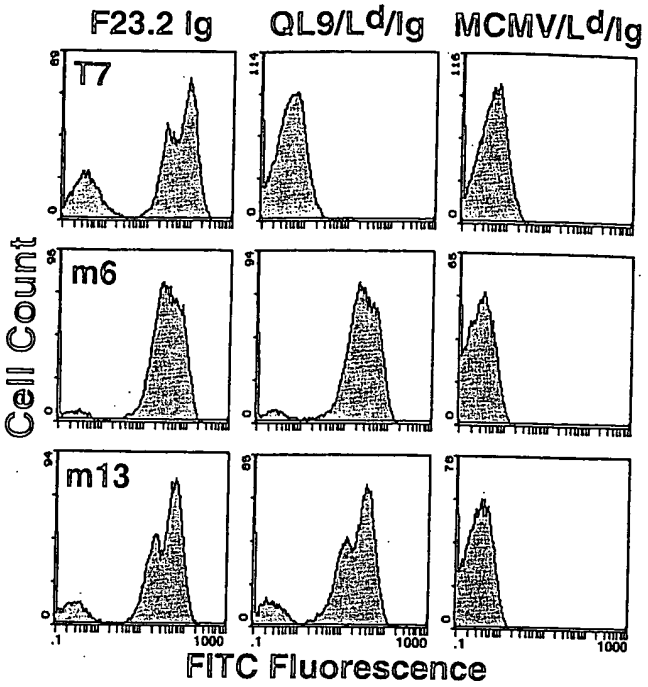
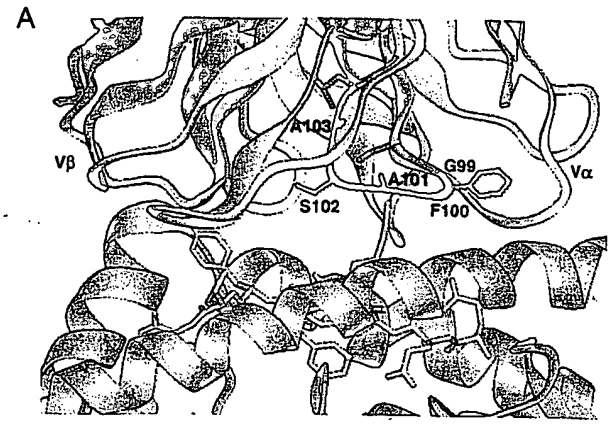


Fig. 1. Flow cytometric analysis of yeast cells that express wild-type and mutant 2C TCR on their surface. Yeast cells displaying wild-type (T7) and mutant (m6 and m13) scTCR were stained with anti-Vβ8 antibody F23.2 (120 nM), the specific alloantigenic peptide-MHC, QL9/L^d/Ig (40 nM), or a null peptide MCMV/L^d/Ig (40 nM). The peptides used in this study were QL9 (QLSPFPFDL), MCMV (YPHFMPTNL), and p2Ca (LSPFPFDL). Binding was detected by FITC-conjugated goat anti-mouse IgG F(ab')$_2$ and analyzed by flow cytometry. The negative population (e.g., seen with F23.2 staining) has been observed for all yeast-displayed proteins and is thought to be caused by cells at a stage of growth or induction that are incapable of expressing surface fusion protein (16, 17, 27).



B

Wild Type TCR	VαCDR3
2C	^{93}SGFASAL^{104}

Mutant TCR	VαCDR3	Mutant TCR	VαCDR3
m1	SSYGNYL	m10	SLPPPLL
m2	SRRGHAL	m11	SIPTPSL
m3	SSRGTAL	m12	SNPPPLL
m4	SHFGTRL	m13	SDPPPLL
m5	SMFGTRL	m14	SSPPPRL
m6	SHQGRYL	m15	SAPPPIL
m7	SYLGLRL		
m8	SKHGIHL		
m9	SLTGRYL		

Fig. 2. Structure and sequences of the 2C TCR CDR3α. (*A*) X-ray crystallographic structure of the 2C/dEV8/K^b complex with CDR3α aa highlighted. Five residues of the 2C VαCDR3 that were randomized by PCR with a degenerate primer are shown in red. The adjacent CDR3 residues, Ser-93 and Leu-104 shown in blue, were retained in the yeast display library because they have been shown to be important in pMHC binding (17, 18, 20). (*B*) Alignment of aa sequences of mutant scTCRs isolated by yeast display and selection with QL9/L^d. Display plasmids were isolated from yeast clones after selection and sequenced to determine CDR3α sequences. Mutants m1, m2, m3, m4, m10, and m11 were isolated after the third round of sorting. All other mutants were isolated after the fourth round of sorting.

indicating that soluble forms of the TCR selected with the yeast display system can serve as probes for tumor-associated pMHC or other T cell-specific ligands.

Materials and Methods

Library Construction. The 2C single-chain TCR (scTCR) used as the scaffold for directed evolution (T7) contained six mutations (βG17E, βG42E, βL81S, αL43P, αW82R, and αI118N) that have been shown to increase the stability of the TCR but still allow pMHC binding (E.V.S., K.D.W., and D.M.K., unpublished results; and ref. 18). Mutagenic PCR of the T7 scTCR VαCDR3 was performed by using an AGA-2-specific upstream primer and a degenerate downstream primer 5′-CTTTTGTGCCGGATC-CAAATGTCAG(SNN)$_5$GCTCACAGCACAGAAGTACACG-GCCGAGTCGCTC-3′. Underlined bases indicate the positions of silent mutations introducing unique *Bam*HI and *Eag*I restriction sites. The purified PCR product was digested with *Nde*I and *Bam*HI and ligated to *Nde*I–*Bam*HI-digested T7/pCT302 (16–18). The ligation mixture was transformed into DH10B electrocompetent *Escherichia coli* (GIBCO/BRL), and transformants were pooled into 250-ml LB supplemented with ampicillin at 100 μg/ml and grown overnight at 37°C. Plasmid DNA was transformed into the yeast strain EBY100 by the method of Gietz and Schiestl (21).

Cell Sorting. The yeast library (22) was grown in 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids (Difco) at 30°C to an OD_{600} = 4.0. To induce surface scTCR expression, yeast were pelleted by centrifugation, resuspended to an OD_{600} = 1.0 in 2% galactose/0.67% yeast nitrogen base/1% Casamino acids, and incubated at 20°C for ≈24 h. In general, ≈10^7 cells per tube were incubated on ice for 1 h with 50 μl of QL9/L^d/IgG dimers (23) diluted in PBS (pH 7.4) supplemented with 0.5 mg/ml BSA. After incubation, cells were washed and labeled for 30 min with FITC-conjugated goat anti-mouse IgG F(ab')$_2$ (Kirkegaard & Perry) in PBS (pH 7.4) supplemented with 0.5 mg/ml BSA. Yeast were then washed and resuspended in PBS (pH 7.4) supplemented with 0.5 mg/ml BSA immediately before sorting. Cells exhibiting the highest fluorescence were isolated by using a Coulter 753 bench fluorescence-activated cell sorter. After isolation, sorted cells were expanded in 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids and induced in 2% galactose/0.67% yeast nitrogen base/1% Casamino acids for subsequent rounds of selection. A total of four sequential sorts were performed. The concentrations of QL9/L^d/IgG dimers used for staining were 50 μg/ml for sorts one to three and 0.5 μg/ml for the final sort. The percentages of total cells isolated from each sort were 5.55, 2.68, 2.56, and 0.58%, respectively. Aliquots of sorts three and four were plated on 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids to isolate individual clones which were analyzed by flow cytometry by using a Coulter Epics XL instrument.

Soluble scTCR Production. The T7 and m6 scTCR genes were excised from pCT302 *Nhe*I–*Xho*I and ligated into *Nhe*I–*Xho*I

digested pRSGALT, a yeast expression plasmid (18). Ligations were transformed into DH10B electrocompetent *E. coli* (GIBCO/BRL). Plasmid DNA was isolated from bacterial cultures and transformed into the *Saccharomyces cerevisiae* strain, BJ5464 (*α ura3–52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL*) (18). Yeast clones were grown in 1 liter of 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids/20 mg/liter Trp for 48 h at 30°C. To induce scTCR secretion, cells were pelleted by centrifugation at 4,000 × *g*, resuspended in 1 liter of 2% galactose/0.67% yeast nitrogen base/1% Casamino acids/20 mg/l Trp supplemented with 1 mg/ml BSA, and incubated for 72 h at 20°C. Culture supernatants were harvested by centrifugation at 4,000 × *g*, concentrated to ≈50 ml, and dialyzed against PBS (pH 8.0). The six His-tagged scTCRs were purified by native nickel affinity chromatography [Ni-NTA Superflow, Qiagen (Chatsworth, CA); 5 mM and 20 mM imidazole (pH 8.0) wash; 250 mM imidazole elution] (18).

Cell-Binding Assays. The binding of soluble scTCRs to QL9/L^d was monitored in a competition format as described (20, 24). Peptide-loaded T2-L^d cells (3×10^5 per well) were incubated for 1 h on ice in the presence of ^{125}I-labeled anti-L^d Fabs (30–5-7) and various concentrations of scTCRs. Bound and unbound [^{125}I] 30–5-7 Fabs were separated by centrifugation through olive oil/dibutyl phthalate. Inhibition curves were constructed to determine inhibitor concentrations yielding 50% maximal inhibition. Dissociation constants were calculated by using the formula of Cheng and Prusoff (25). To monitor direct binding of scTCRs to cell-bound pMHC, peptide-loaded T2-L^d cells (5×10^5 per tube) were incubated for 40 min on ice with biotinylated soluble scTCRs followed by staining for 30 min with streptavidin-phycoerythrin (PharMingen). Cellular fluorescence was detected by flow cytometry.

Results and Discussion

To examine if it is possible to generate higher-affinity TCR that would retain peptide specificity, we subjected a TCR to a process of directed *in vitro* evolution. Phage display (26) has not yet proven successful in the engineering of single-chain TCRs (scTCRs, Vβ-linker-Vα) despite the extensive structural similarity between antibody and TCR V regions. However, we recently showed that a scTCR could be displayed on the surface of yeast (17) in a system that has proven successful in antibody engineering (16, 27). A temperature-stabilized variant (T7) (18) of the scTCR from the cytotoxic T lymphocyte clone 2C was used in the present study. Cytotoxic T lymphocyte clone 2C recognizes the alloantigen L^d with a bound octamer peptide called p2Ca, derived from the enzyme 2-oxoglutarate dehydrogenase (28). The nonameric variant QL9 is also recognized by cytotoxic T lymphocyte 2C, but with 10-fold higher affinity by the 2C TCR (29). Alanine scanning mutagenesis showed that the CDR3α loop contributed minimal free energy to the binding interaction (20), even though structural studies have shown that CDR3α of the 2C TCR is near the peptide and it undergoes a conformational change to accommodate the pMHC complex (19). Thus, we focused our mutagenesis efforts on five residues that form the CDR3α loop.

A library of 10^5 independent TCR-CDR3α yeast mutants was subjected to selection by flow cytometry with a fluorescently labeled QL9/L^d ligand (23). After four rounds of sorting and growth, 15 different yeast colonies were examined for their ability to bind the ligand, in comparison to the scTCR variant T7, which bears the wild-type CDR3α sequence (Fig. 1 and data not shown). The anti-Vβ8.2 antibody F23.2 which recognizes residues in the CDR1 and CDR2 was used as a control to show that wild-type scTCR-T7 and scTCR mutants (m6 and m13, in Fig. 1, and others, data not shown) each had approximately equivalent surface levels of the scTCR (Fig. 1). In contrast, the soluble QL9/L^d ligand bound very well to each mutant yeast clone but not to wild-type scTCR-T7. The MCMV/L^d complex, which is not recognized by cytotoxic T lymphocyte clone 2C, did not bind

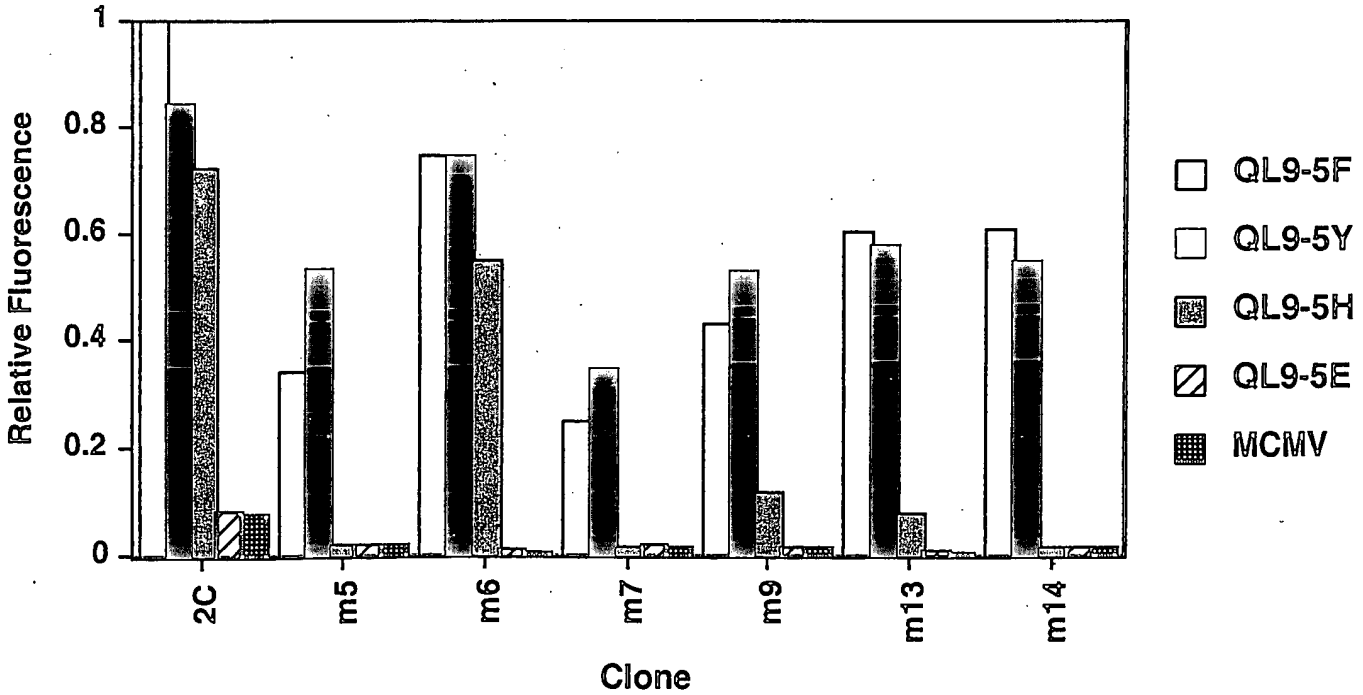


Fig. 3. Fine specificity analysis of mutant scTCR binding to different QL9 variant peptides bound to L^d. The original T cell clone 2C and various yeast clones were analyzed by flow cytometry for binding to L^d/Ig dimers loaded with wild-type QL9 (QL9–5F), position 5 variants of QL9 (QL9–5Y, QL9–5H, and QL9–5E) or MCMV. Binding was detected with FITC-labeled goat anti-mouse IgG. Relative fluorescence was measured by comparison with mean fluorescence values of 2C cells or yeast cells stained with anti-Vβ8 antibody F23.2.

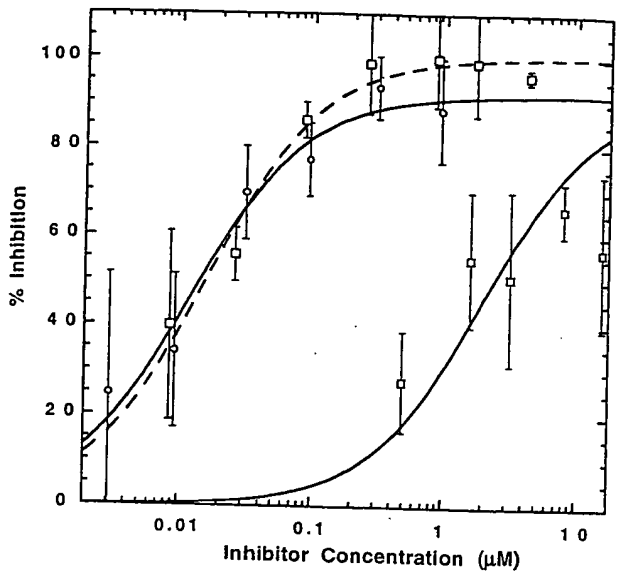


Fig. 4. QL9/L^d binding by soluble scTCRs. T2-L^d cells loaded with QL9 were incubated with ^{125}I-labeled anti-L^d Fab fragments (30-5-7) and various concentrations of unlabeled Fab (□), scTCR-T7 (⊡), or mutant scTCR-m6 (⊙). Bound and unbound [^{125}I]30-5-7 Fab fragments were separated by centrifugation through olive oil/dibutyl phthalate. Binding of ^{125}I-labeled anti-L^d Fab fragments to T2-L^d cells loaded with the control peptide MCMV was not inhibited even at the highest concentrations of scTCRs (data not shown).

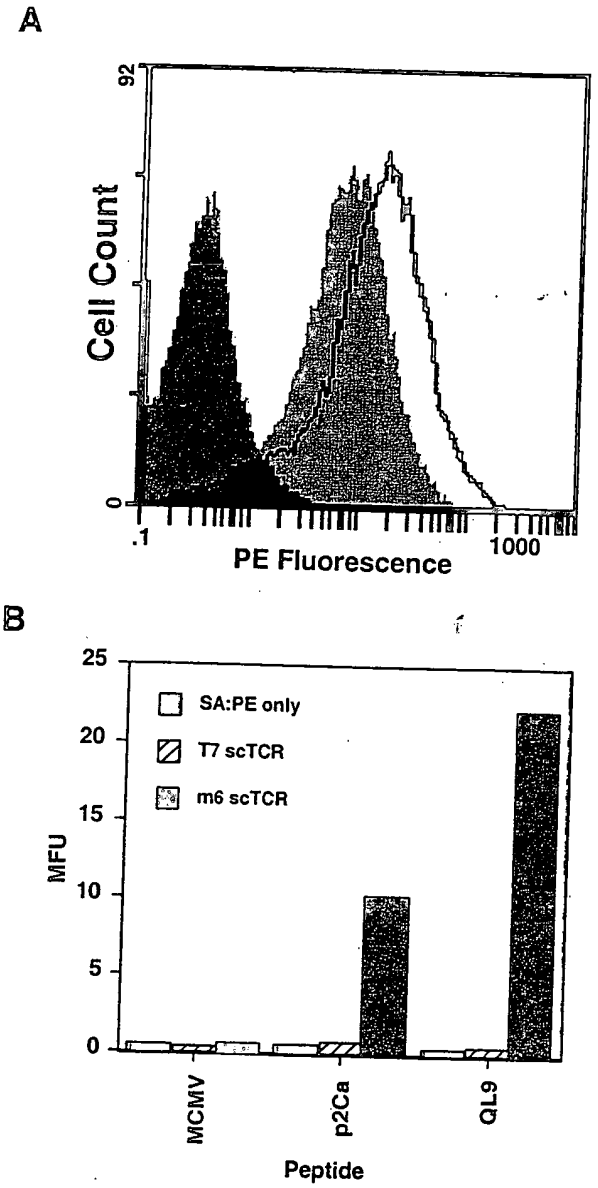


Fig. 5. Flow cytometric analysis of the binding of scTCR/biotin to cell surface peptide/MHC. Peptide-loaded T2-L^d cells were incubated with biotinylated m6 scTCR (~0.3 μM) or T7 scTCR (~1.6 μM) scTCR followed by streptavidin-PE and analyzed by flow cytometry. (*A*) Flow cytometry histograms of T2-L^d cells loaded with QL9 (unshaded), p2Ca (light shade), or MCMV (dark shade) and stained with m6 scTCR/biotin. (*B*) Mean fluorescent units (MFU) of T2-L^d cells loaded with QL9, p2Ca, or MCMV and stained with either secondary SA-PE only, T7 scTCR/biotin + SA-PE, or m6 scTCR/biotin + SA-PE.

to the scTCR mutants or wild-type scTCR-T7, indicating that the scTCR mutants retained peptide specificity.

The CDR3α sequences of the 15 mutants all differed from the 2C TCR (Fig. 2). It was readily apparent (and confirmed by a BLAST alignment algorithm) that the sequences could be aligned into two motifs. One motif contained Gly in the middle of the five residue stretch whereas the other motif contained three tandem Pro. Evidence that all three Pro are important in generating the highest affinity site is suggested by results with mutant m11. Mutant m11 contained only two of the three Pro and exhibited reduced binding compared with the triple-Pro mutants (data not shown). The Gly-containing mutants appeared to have preferences for positive-charged residues among the two residues to the carboxyl side (7/9) and aromatic and/or positive-charged residues among the two residues to the amino side (4/9 and 5/9). The selection for a glycine residue at position 102 in the motif may indicate that the CDR3α loop required conformational flexibility around this residue to achieve increased affinity. This is consistent with the large (6-Å) conformational difference observed between the CDR3α loops of the liganded and unliganded 2C TCR (19). It is also interesting that Gly is the most common residue at the V(D)J junctions of antibodies and the presence of a Gly has recently been associated with increased affinity in the response to the (4-hydroxy-3-nitrophenyl) acetyl hapten (30).

In contrast to the isolates that contain Gly, the selection for a Pro-rich sequence at the tip of the CDR3α loop may suggest that these TCR exhibit a more rigid conformation that confers higher affinity. The x-ray crystallographic structures of a germline antibody of low affinity compared with its affinity-matured derivative showed that the high-affinity state was associated with stabilization of the antibody in a configuration that accommodated the hapten (31). Similarly, the NMR solution structure of a scTCR that may be analogous to the germ-line antibody showed that the CDR3α and β loops both exhibited significant mobility (32). Recent thermodynamic studies of TCR:pMHC interactions have also suggested the importance of conformational changes in binding (33, 34). Structural and thermodynamic studies of the TCR mutants that we report here should allow us to examine if the two CDR3α motifs (Gly- versus Pro-rich) might differ in the mechanism by which they confer higher affinity.

Although the scTCR mutants did not bind the null peptide/L^d complex MCMV/L^d, it remained possible that the increase in affinity might be accompanied by a change in fine specificity. To examine this issue, we used QL9 position 5 (Phe) peptide variants that have been shown previously to exhibit significant differences in their binding affinity for the wild-type 2C TCR (35). The binding of these pMHC to various TCR mutants on the yeast surface and clone 2C were measured by flow cytometry. As

shown in Fig. 3, the native TCR on 2C is capable of binding QL9 variants that contain either Tyr or His at position five but not Glu. Each of the higher-affinity TCR mutants retained their ability to recognize the conserved Tyr-substituted peptide and they were likewise incapable of recognizing the Glu-substituted peptide. However, several of the TCR mutants (m6, m9, and m13) bound to the His-substituted peptide (albeit to different extents) whereas other mutants (m5, m7, and m14) did not bind the peptide. Thus, the CDR3α loop can influence the peptide fine specificity of recognition but it is not the only region of the TCR involved. The effect on peptide specificity could be through direct interaction of CDR3α residues with the variant peptide, as suggested from earlier studies involving CDR3-directed selections (36, 37). Alternatively, binding free energy may be directed at peptide-induced changes in the L^d molecule itself. The latter possibility is perhaps more likely in the case of the 2C TCR:QL9/L^d interaction, because position five of QL9 has been predicted to point toward the L^d groove (35, 38). The fine-specificity analysis also shows that it is possible to engineer TCRs with increased, or at least altered, specificity for cognate peptides. Thus, directed evolution of only a short region (CDR3α) of a single TCR allows the design of TCR variants with altered peptide-binding specificities.

To determine the magnitude of the affinity increases associated with a selected CDR3α mutant, the wild-type T7 scTCR and the m6 scTCR were expressed as soluble forms in a yeast secretion system. Purified scTCR preparations were compared for their ability to block the binding of a ^{125}I-labeled anti-L^d Fab fragments to QL9 or MCMV loaded onto L^d on the surface of T2-L^d cells. As expected, neither T7 nor m6 scTCR were capable of inhibiting the binding of ^{125}I-Fab fragments to T2-L^d cells up-regulated with the MCMV peptide (data not shown). However, both T7 and m6 were capable of inhibiting the binding of anti-L^d Fab fragments to QL9/L^d (Fig. 4). The m6 scTCR variant was as effective as unlabeled Fab fragments in inhibiting binding, whereas the T7 scTCR was 160-fold less effective (average of 140-fold difference among four independent titrations). The K_D values of the scTCR for QL9/L^d were calculated from the inhibition curves to be 1.5 μM for T7 and 9.0 nM for m6. The value for T7 is in close agreement with the 3.2 μM K_D reported for the 2C scTCR (39). These findings show that the yeast system, combined with CDR3α-directed mutagenesis, is capable of selecting mutants with 100-fold higher intrinsic binding affinities for a pMHC ligand.

If the soluble scTCR has a high affinity for its pMHC ligand, then it may be useful, like antibodies, as a specific probe for cell surface-bound antigen. To test this possibility, the soluble T7 and m6 scTCR were biotinylated and the labeled scTCR were incubated with T2-L^d cells loaded with QL9, p2Ca, or MCMV. The m6 scTCR, but not the T7 scTCR, yielded easily detectable staining of the T2 cells that had been incubated with QL9 or p2Ca (Fig. 5 *A* and *B*). It is significant that p2Ca-up-regulated cells were also readily detected by m6 scTCR, because p2Ca is the naturally processed form of the peptide recognized by the alloreactive clone 2C and it has an even lower affinity than the QL9/L^d complex for the 2C TCR (29). However, it remains to be determined if the levels of pMHC derived from endogenous antigen processing are sufficient to allow detection by using soluble TCR as probes. It is reasonable to predict that, in some cases, the level will be too low to distinguish from background by using standard flow cytometry procedures.

The high-affinity receptors described in our study were derived by variation at the VJ junction, the same process that operates very effectively *in vivo* through gene rearrangements in T cells (2). The fact that we could readily isolate a diverse set of high-affinity TCR *in vitro* indicates that there is not a genetic or structural limitation to high-affinity receptors. This supports the view that inherently low affinities of TCRs found *in vivo* are caused by a lack of selection for higher affinity and perhaps a selection for lower affinity (5–7). In this respect, the higher-affinity TCRs now provide the reagents for directly testing hypotheses about the effects of affinity on T cell responses (4–7). It is interesting to note that similar arguments have been used to suggest that the kinetic properties of antibodies may also set an *in vivo* "affinity ceiling," above which there may not be a selective advantage to B cells (40).

In addition to their utility for testing T cell responses, high-affinity TCRs can be engineered like antibodies to yield high-affinity, antigen-specific probes. Soluble versions of the high-affinity receptor can directly detect specific peptide/MHC complexes on cells (Fig. 5). Thus, these engineered proteins have potential, for example, as tumor cell diagnostics, or on conjugation with cytotoxins, potential agents for cancer therapy.

We thank Michele Kieke for providing 2C scTCR genes cloned in the yeast display vector and Gary Durack of the University of Illinois Biotechnology Center Flow Cytometry Facility for assistance and advice. We also thank Jeff Bluestone and Ian Wilson for comments and discussion on the manuscript and results. This work was supported by National Institutes of Health Grant RO1 GM55767.

1. Tonegawa, S. (1988) *Biosci. Rep.* **8**, 3–26.
2. Davis, M. M. & Bjorkman, P. J. (1988) *Nature (London)* **334**, 395–402.
3. Eisen, H. N., Sykulev, Y. & Tsomides, T. J. (1996) *Adv. Protein Chem.* **49**, 1–56.
4. Davis, M. M., Boniface, J. J., Reich, Z., Lyons, D., Hampl, J., Arden, B. & Chien, Y. (1998) *Annu. Rev. Immunol.* **16**, 523–544.
5. Sykulev, Y., Cohen, R. J. & Eisen, H. N. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 11990–11992.
6. Valitutti, S., Muller, S., Cella, M., Padovan, E. & Lanzavecchia, A. (1995) *Nature (London)* **375**, 148–151.
7. Rabinowitz, J. D., Beeson, C., Lyons, D. S., Davis, M. M. & McConnell, H. M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 1401–1405.
8. Rabinowitz, J. D., Beeson, C., Wulfing, C., Tate, K., Allen, P. M., Davis, M. M. & McConnnell, H. M. (1996) *Immunity* **5**, 125–135.
9. Sloan-Lancaster, J. & Allen, P. M. (1996) *Annu. Rev. Immunol.* **14**, 1–27.
10. Alam, S. M., Travers, P. J., Wung, J. L., Nasholds, W., Redpath, S., Jameson, S. C. & Gascoigne, R. J. (1996) *Nature (London)* **381**, 616–620.
11. Kersh, G. J., Kersh, E. N., Fremont, D. H. & Allen, P. M. (1998) *Immunity* **9**, 817–826.
12. Al-Ramadi, B. K., Jelonek, M. T., Boyd, L. F., Margulies, D. H. & Bothwell, A. L. M. (1995) *J. Immunol.* **155**, 662–673.
13. Sykulev, Y., Vugmeyster, Y., Brunmark, A., Ploegh, H. L. & Eisen, H. N. (1998) *Immunity* **9**, 475–483.
14. Bjorkman, P. J. (1997) *Cell* **89**, 167–170.
15. Bevan, M. J. (1997) *Immunity* **7**, 175–178.
16. Boder, E. T. & Wittrup, K. D. (1997) *Nat. Biotechnol.* **15**, 553–557.
17. Kieke, M. C., Shusta, E. V., Boder, E. T., Teyton, L., Wittrup, K. D. & Kranz, D. M. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 5651–5656.
18. Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M. & Wittrup, K. D. (1999) *J. Mol. Biol.* **292**, 949–956.
19. Garcia, K. C., Degano, M., Pease, L. R., Huang, M., Peterson, P., Teyton, L. & Wilson, I. A. (1998) *Science* **279**, 1166–1172.
20. Manning, T. C., Schlueter, C. J., Brodnicki, T. C., Parke, E. A., Speir, J. A., Garcia, K. C., Teyton, L., Wilson, I. A. & Kranz, D. M. (1998) *Immunity* **8**, 413–425.
21. Gietz, R. D., Schiestl, R. H., Willems, A. & Woods, R. A. (1995) *Yeast* **11**, 355–360.
22. Shusta, E. V., VanAntwerp, J. & Wittrup, K. D. (1999) *Curr. Opin. Biotechnol.* **10**, 117–122.
23. Dal Porto, J., Johansen, T. E., Catipovic, B., Parfiit, D. J., Tuveson, D., Gether, U., Kozlowski, S., Fearon, D. T. & Schneck, J. P. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6671–6675.
24. Sykulev, Y., Brunmark, A., Jackson, M., Cohen, R. J., Peterson, P. A. & Eisen, H. N. (1994) *Immunity* **1**, 15–22.
25. Cheng, Y. C. & Prusoff, W. H. (1973) *Biochem. Pharmacol.* **22**, 3099–3108.
26. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991) *Nature (London)* **352**, 624–628.
27. Kieke, M. C., Cho, B. K., Boder, E. T., Kranz, D. M. & Wittrup, K. D. (1997) *Protein Eng.* **10**, 1303–1310.
28. Udaka, K., Tsomides, T. J., Walden, P. P., Fukusen, N. & Eisen, H. N. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11272–11276.

29. Sykulev, Y., Brunmark, A., Tsomides, T. J., Kageyama, S., Jackson, M., Peterson, P. A. & Eisen, H. N. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 11487–11491.
30. Furukawa, K., Akasako-Furukawa, A., Shirai, H., Nakamura, H. & Azuma, T. (1999) *Immunity* **11**, 329–338.
31. Wedemayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G. & Stevens, R. C. (1997) *Science* **276**, 1665–1669.
32. Hare, B. J., Wyss, D. F., Osburne, M. S., Kern, P. S., Reinherz, E. L. & Wagner, G. (1999) *Nat. Struct. Biol.* **6**, 574–581.
33. Willcox, B. E., Gao, G. F., Wyer, J. R., Ladbury, J. E., Bell, J. I., Jakobsen, B. K. & Anton van der Merwe, P. (1999) *Immunity* **10**, 357–365.
34. Boniface, J. J., Reich, Z., Lyons, D. S. & Davis, M. M. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 11446–11451.
35. Schlueter, C. J., Manning, T. C., Schodin, B. A. & Kranz, D. M. (1996) *J. Immunol.* **157**, 4478–4485.
36. Sant'Angelo, D. B., Waterbury, G., Preston-Hurburt, P., Yoon, S. T., Medzhitov, R., Hong, S. & Janeway, C. A. (1996) *Immunity* **4**, 367–376.
37. Jorgensen, J. L., Esser, U., de St. Groth, B. F., Reay, P. A. & Davis, M. M. (1992) *Nature (London)* **355**, 224–230.
38. Speir, J. A., Garcia, K. C., Brunmark, A., Degano, M., Peterson, P. A., Teyton, L. & Wilson, I. A. (1998) *Immunity* **8**, 553–562.
39. Manning, T. C., Parke, E. A., Teyton, L. & Kranz, D. M. (1999) *J. Exp. Med.* **189**, 461–470.
40. Foote, J. & Eisen, H. N. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1254–1256.

Docket No. 46745 (48340)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

APPLICANT:	J. Weidanz, et al.		
SERIAL NO.:	08/813,781	GROUP:	1644
FILED:	March 7, 1997	EXAMINER:	R. Schwadron

FOR: FUSION PROTEINS COMPRISING BACTERIOPHAGE COAT PROTEIN AND A SINGLE-CHAIN T CELL RECEPTOR

THE HONORABLE COMMISSIONER OF PATENTS AND TRADEMARKS
WASHINGTON, D.C. 20231

APPELLANTS' BRIEF ON APPEAL
SUBMITTED PURSUANT TO 37 C.F.R. §1.192

In support of Appellants' appeal on October 17, 2002 of the Examiner's final rejection, mailed on June 17, 2002, submitted herewith is Appellants' Brief on Appeal.

Table of Contents

REAL PARTY IN INTEREST

The real party in interest is Altor BioScience Corporation of Miramar, Florida. An assignment from the inventors to Dade International was recorded on August 18, 1997 at Reel/ Frame 8681/0081. An assignment from Dade International to Sunol Molecular Corporation was recorded on February 26, 2003 at Reel/Frame 013787/0276. An assignment from Sunol Molecular Corporation to Altor BioScience Corporation was mailed to the USPTO on February 3, 2004.

RELATED APPEALS AND INTERFERENCES

There are no related appeals or interferences known to Appellants or Appellants' representatives that will directly affect or be directly affected by or have a bearing on the Board's decision in the pending Appeal.

STATUS OF THE CLAIMS

Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 stand finally rejected under 35 U.S.C. §103 and are pending on appeal. Claims 5, 6, 10-12, 16-20, 61, 65, 70, 73 and 74 have been cancelled. Claims 3, 9, 13, 15, 21-60, 62-64, 66 and 68 were withdrawn from consideration by the Examiner.

STATUS OF THE AMENDMENTS

Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 on appeal are set forth in Exhibit A hereto. There has been no further amendment to the claims.

SUMMARY OF THE CLAIMED INVENTION

The claimed invention features a soluble fusion protein engineered to include a bacteriophage coat protein fused to a single-chain T cell receptor ("scTCR"). The single-chain T cell receptor was itself designed to include an alpha-variable region ("V-α") fused to a beta-variable region ("V-β"). The single-chain T cell receptor forms a pocket that binds antigen when the antigen. The claimed soluble fusion protein further includes a beta-constant region ("C-β") region that can be fused to V-β, for example.

T cells help defend the body against infection. The cells have membrane bound receptors that bind foreign antigen with the assistance of a protein complex called "MHC". A key receptor is called the T cell receptor ("TCR"). The chemical structure and function of the TCR has been extensively studied. For instance, it is known that formation of a TCR-antigen-MHC complex is an important step toward fighting infection.

Appellants discovered that by adding a bacteriophage coat protein to the scTCR, it is possible to produce a fully soluble and functional scTCR. Unlike prior scTCRs, the claimed fusion proteins were found to be fully soluble, functional, and obtainable in significant quantities without difficulty. The claimed fusion proteins have a wide spectrum of important uses as described throughout the instant patent application.

See the Summary Of The Invention at pg. 3, line 22 to pg. 11, line 21. See also the Background at pg. 1, line 12 to pg. 3, line 19 for related information. None of the art of record in this case shows an attempt to make a scTCR that includes a fused bacteriophage coat protein.

ISSUE

This appeal presents the issue of whether the Examiner erred in rejecting claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 under 35 U.S.C. § 103 in view of Chung, S. et al. (1994) *Proc. Natl. Acad. Sci. (USA)* 91: 12654 in view of U.S Pat. No. 5,759,817 to Barbas, Onda, T et al. (1996) *Mol. Immunol.* 32: 1387; and Huse et al. (1992) *J. Immunol.* 149: 3914. Appellants will refer to these citations as "Chung", "Barbas", "Onda" and "Huse", respectively; unless stated otherwise. There are no other pending rejections of record in this case.

GROUPING OF THE CLAIMS

All of claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 stand or fall together for the purpose of the present appeal.

CASE HISTORY

In consideration of the Examiner's position in this case, Appellants have summarized the prosecution history with respect to the 35 U.S.C. §103 rejection at issue. A more detailed discussion of Appellants' rebuttal to those arguments will follow under Argument.

A. A patent application was filed on March 7, 1997 with 59 claims and assigned Serial No.: 08/813,781 by the USPTO.

B. An Office Action was mailed to Appellants on August 25, 2000 by Examiner Schwadron. The Examiner stated the following with respect to the §103 rejection on appeal:

> Claims 1, 2, 4, 7, 8, 14, 67, 69, 71, 72 are rejected under 35 U.S.C. 103(a) as unpatentable over Chung et al. in view of Barbas US 5,759,817 (filed Jan. 27, 1992), Onda et al. (Molecular Immunology 32:1387, 1995), and Huse et al. J. Immunology 149:3914. 1992
>
> Chung et al. teaches a single chain T cell receptor which specifically binds to peptide ligand (see abstract). Chung et al. further teaches one embodiment of human single chain TCR in which C-terminus of V α domain is linked to N-terminus of V β chain via a 15 amino acid residue flexible amino acid liner and the C-terminus of the V β chain is linked to the beta chain constant domain (see Figure 1). In one embodiment the C terminus of V β chain is linked to a alkaline phosphatase (PI) protein tag (see page 12655). Chung et also teach that the purpose of the linker is to enhance the binding characteristics of the soluble T cell receptor and that linkers of about 10 to 30 amino acid residues would be considered to be sufficient. Chung et al. teach that the TCR fusion protein can bind antigenic protein, thus teaching that the TCR fusion protein comprises an antigen binding pocket. Chung et al. teaches a TCR fusion protein comprising V-α-peptide linker-V β-C β linked to GPI anchor and expression of such a fusion protein in a transfected eukaryotic cell (see results section). Chung et al. disclose that the soluble form of TCR protein could be readily obtained by enzymatic cleavage with phosphatidylinostol-specific phospholipase C (PI-PLC) (see page 12656). Chung et al. teaches expression of said TCR fusion protein in a bacterial cell system in which the N terminus of the Cβ region is linked to a histidine protein tag. Chung et al. also disclose a scTCR in which comprises V-α-peptide linker-V β-C β GPI in which the Cβ component consists of the β chain sequence ending right before the last cysteine (the sixth cysteine) (see page 12655). Chung et al. further teach that TCR fusion proteins which do not contain the CB do not fold into the native conformation. The scTCR disclosed by Chung et al. meet the length limitations of the Vα. and Vβ region recited in claims 69 and 71. Chung et al. teach a soluble fusion protein comprising a Vα-peptide linker-Vβ-Cβ fragment-protein tag (eg. GPI). **Chung et al. does not teach a TCR fusion protein further comprising bacteriophage VIII coat protein.**

However, Barbas discloses a soluble fusion protein comprising a bacteriophage coat protein fragment covalently linked to a single-chain heterodimeric receptor (see abstract and column 15, lines 27-28, in particular). Barbas also discloses that the fusion protein may comprise domains of heterodimeric proteins derived from several ligand binding proteins, including immunoglobulins and T cell receptors (see column 17, lines 62-66 and column 19, lines, 9-28. Barbas discloses that T cell receptor comprises alpha and beta chains each having a variable(V) and constant(C) region and T cell receptor has similarities in genetic organization and function to immunoglobulins (see column 19, lines 19-22, in particular). Barbas also teaches that bacteriophage coat protein may be derived from cpIII or cpVIII (see column 31, lines 10-28, in particular). Barbas discloses that expression vectors expressing soluble fusion proteins in which the ligand binding region is fused to bacteria coat protein allows the expression of the multiple fusion proteins on the surface of phage particles IE approximately 2700 cpVIII heterodimer receptor molecules per phage particle (see column 39 line 64 through column 40, line 7, in particular). Barbas further discloses that a short length of amino acid sequence at the amino end of a protein (IE a protein tag) directs the protein to periplasmic space (see column 8, lines 49-55, in particular. One embodiment of the invention is disclosed to be a fusion protein comprising in sequence a leader sequence-peptide linker-V region amino acid residue-peptide linker-phage coat protein and that in one embodiment, the second linker can define a proteolytic cleavage site which allows the heterodimeric receptor to be cleaved from the bacteriophage coat protein to which it is attached (see column 14, lines 60-65). **Thus Barbas discloses but does not exemplify a soluble fusion protein comprising a bacteriophage coat protein covalently linked to T cell receptor domains.**[1]

Onda et al. disclose a soluble fusion protein comprising a bacteriophage coat protein covalently linked to a single-chain T cell receptor by a peptide linker sequence wherein the single TCR chain is the alpha chain and the bacteriophage coat protein is cpVIII (see abstract and Figure 1, in particular). Onda et al. also teach that TCR-bacteriophage coat protein fusion protein can be used to study specific binding interactions of the TCR chain to antigenic ligands (see paragraph bridging pages 1394-1395, in particular).

Huse et al. teach that fusion proteins comprising a single chain fusion protein comprising Fab fragment of immunoglobulin (which comprises the antigen binding pocket of the immunoglobulin molecule) and bacteriophage VIII coat protein can be produced and display the fusion protein when expressed in a M13 derived vector. Huse et al. further teach that bacteriophage VIII coat protein

[1] Examiner Schwadron withdrew this statement in the next Office Action dated June 17, 2002 (see below). The statement was originally made by a prior Examiner (Lubet) in a related §103 rejection that has been withdrawn. In that earlier rejection, Lubet argued that Barbas and Onda **do not teach** a soluble fusion protein in which a single-chain TCR linked to a bacteriophage coat protein. See the Office Action dated June 23, 1998 at pg. 8, part B. In the §103 rejection on appeal, Examiner Schwadron argued that Barbas and Onda teach use of TCR-bacteriophage VIII fusion protein. That molecule, a heterodimeric T cell receptor fusion protein, is not the claimed invention.

fusion protein can recovered from culture medium or from the periplasmic space (see abstract).

Therefore it would have been prima facie obvious to one with ordinary skill in the art at the time the invention was made to make a soluble TCR fusion protein comprising the Vα..-peptide linker-Vβ -Cb fragment-protein taught by Chung et al. linked to a bacteriophage VIII coat protein because **Barbas et al. and Onda et al. teach TCR-bacteriophage VIII coat fusion proteins can be used to study antigen binding properties of such a fusion protein** and Huse et al. teach that fusion proteins comprising bacteriophage VIII coat protein can be produced in bacteria and recovered in relatively large quantities.

One with skill in the art would be motivated to make such a fusion protein to study the antigen binding region of the TCR component or to use the protein to elicit anti-idiotypic antibodies. One with skill in the art would be motivated to make such a fusion protein in which the Vα and Vβ region was derived from human TCR in order to study human TCR properties or to elicit anti-idiotypic antibodies to the TCR component of the protein.

The preceding grounds of rejection have been maintained since the August 25, 2000 Office Action despite Appellants' rebuttal argument and claim amendments discussed below.

C. On October 25, 2000, Appellants' representative met with Examiner Schwadron at the USPTO and discussed the art cited. No agreement was reached.

D. Appellants submitted a response to the rejection set forth in paragraph B, above, on February 22, 2001. In that response, Appellants rebutted the prima facie obviousness argument by pointing out: 1) that the cited references did not teach or suggest that the membrane "anchor" of Chung's single-chain TCR (GPI: a membrane protein) could be substituted with the bacteriophage coat protein of Barbas' TCR; 2) that there was no reasonable expectation that the substitution (switching Chung's anchor for Barbas' phage coat protein) could be achieved in view of substantial differences between scTCRs and TCRs; and 3) that the Examiner's citation of Onda was not correct ie., it does not disclose TCR-bacteriophage coat protein fusions, but instead, dwelt on smaller constructs having only a V-α chain (but no V-β chain). Onda characterized his constructs as having "unusual" binding properties that were not characteristic of TCRs. A subset of such constructs were reported by Onda not to work at all.

Appellants also discussed the **Holler** reference: a peer-reviewed scientific article from the U.S. Academy of Sciences (*PNAS (USA)* (2000) 97: 5387 at 5389). Holler provided independent and objective evidence of the **long-felt need and failure of others** in the field to make and use the claimed fusion molecules. Specifically, Holler stated that phage display had not yet proven successful in the engineering of scTCRs.[2]

E. In response to the Appellants' arguments in paragraph D above, the Examiner issued a Final Office Action dated June 17, 2002.[3] The Examiner maintained the prima facie rejection and stated:

> Regarding appellants comments, while **heterodimeric molecules are a preferred embodiment** disclosed in Barbas et al., Barbas et al. disclose: "In another embodiment, the present invention contemplates a polypeptide comprising an insert domain flanked by an amini-terminal secretion signal domain and a carboxy-terminal filamentous phage coat protein membrane anchor domain." (column 14, first complete paragraph).
>
> Barbas et al. further disclose than said construct could include a **"receptor protein"** (column 14, second paragraph), indicating that the disclosed method could be used for receptors per se (eg. single chain or heterodimeric or single chain heteromers). Single chain T cell receptors were known in the art (see Chung et al.).
>
> Regarding appellants comments about the single chain TCR taught by Chung et al.,
>
> Chung et al. teach that the GPI anchor is cleaved and the soluble TCR still has all the antigen binding properties of the TCR (see pages 12656-12658). Thus, the GPI anchor is not required for the soluble TCR to function, it is just used in one particular method of making the soluble TCR. Regarding motivation to create the claimed invention, **Chung et al. discloses that it would be desirable to produce their TCR in a phage display system (see page 12658, first column).** In addition, Barbas et al. teach the advantages of their system for the production

---

[2] By "phage display" is meant the process of making a recombinant bacteriophage expressing the scTCR as part of the phage protein coat. After infecting bacteria with a recombinant phage engineered to produce the scTCR, the protein would be "displayed" on the bacterial cell surface as bacteriophage. The scTCR "displayed" in this manner would be amendable to engineering.

[3] The Examiner essentially repeated his rejection of the claims as set forth in the August 25, 2000 rejection. But see footnote 1. With respect to Examiner Schwadron's discussion about Holler, Appellants submitted Weidanz et al. as part of a Rule 132 Declaration to address an obviousness rejection that has since been withdrawn. The reference provided evidence that a particular scTCR-bacteriophage coat protein vector (pKC44) was capable of forming an antigen binding site when expressed. The reference is not prior art to the present application.

of peptides. Regarding reasonable expectation of success, both Barbas et al. and Chung et al. disclose use of phage display systems to produce single chain antibodies (see column 2, third paragraph from bottom and page 12658, first column). In addition, the soluble single chain TCR molecules functions with or without the GPI linker indicating that the construct itself is functional.

Regarding appellants comments about Holler et al., said publication was published in May 2000. In the amendment filed 6/3/2000, applicant submitted a publication by Weidanz et al. (J. Imm. Methods 1998) which discloses the claimed invention. Thus, **it appears that Holler et al. simply are not familiar with the prior art. Thus, the comments of Holler et al. carry no weight because two years prior to the Holler et al. publication, Weidanz et al. had already published data regarding the production of single chain TCR using bacteriophage**. Furthermore, Holler et al. discloses a yeast system for producing a single chain TCR and it appears that the main focus of Holler et al. is to promote their system.

Regarding appellants comments about Onda et al., the instant rejection indicates that "**Onda et al. disclose a soluble fusion protein comprising a bacteriophage coat protein covalently linked to a single-chain T cell receptor** by a peptide linker sequence wherein the single TCR chain is the alpha chain and the bacteriophage coat protein is cpVIII (see abstract and Figure 1, in particular)". The art recognizes that the alpha and beta chains of the TCR generally both are involved in antigen binding. The art also recognizes that soluble TCR which bind antigen would have a variety of uses.

F. Appellants filed a Notice of Appeal on October 17, 2002.

ARGUMENTS

As an initial mater, Appellants wish to emphasize the substantial differences between TCR heterodimers ("TCRs") and single-claim T cell receptors ("sc-TCRs").

The TCR is a heterodimer with one α chain and one β chain.[4] Each of these chains passes from the exterior of the T cell, through the cell membrane, and into the cell interior (cytosol). The α and β chains each include a variable (V-α, V-β) region that cooperate to form an antigen binding pocket. The regions are "variable" because its chemical structure can be

[4] A textbook in the field describes the TCR as "a heterodimer composed of an α and a β polypeptide chain, both of which are glycosylated." See Alberts, B et al. (1989) in *Molecular Biology of the Cell*, 2nd Ed. Garland Publishing, Inc. New York at pg. 1037. By convention, a "heterodimer" such as the TCR properly has two chains (dimer) both of which are different (hetero) from the other.

changed to make a pocket that fits another antigen. Each of the V-α and V-β regions are associated with a constant (C) region.

TCR heterodimers have been extremely difficult to isolate from T cells. This problem has hindered study of the receptor. One approach to address the problem has been to make single-chain T cell receptors ("scTCRs"). These synthetic receptors include, on one chain instead of two, a fused V-α and V-β region. It has been customary to space the V-α and V-β regions from each other with a flexible linker to allow the regions to make an antigen binding pocket. Unfortunately, many scTCRs have still proven to be difficult to make and use.

Appellants point out that "TCR" is understood in the field to mean a **heterodimeric** T cell receptor. The TCR is a membrane bound (insoluble) receptor in which the α and β chains cooperate to bind antigen. Reference to a "scTCR" is understood to mean a synthetic **single-chain** molecule that includes the V-α and V-β regions bound together usually through a flexible linker. Unlike the TCR, the scTCR binds antigen with only one chain. The TCR and scTCR are structurally distinct proteins that are different molecules that bind antigen in different ways.

I. Summary of the Cited Art

A. Chung reports functional three-domain single-chain T cell receptors consisting of a human Vα and Vβ region that recognizes a particular antigen (HLA-DR2b/myelin basic protein). Chung determined that it was important to fuse a Cβ region to the Vβ region. Such a three domain construct, when linked to a synthetic cell membrane anchor (glycosyl phosphatidyl-inositol (GPI) or CD3ζ fragment), was found to be expressed and functional. Chung disclosed that the cell membrane anchor could be cleaved from the single-chain receptors to obtain soluble protein. See the Abstract.

Chung opined that his single-chain design "may allow" construction of TCR phage libraries and that such libraries "may be" tools for studying TCRs. See pg. 12658. However, there is no specific disclosure in Chung about how such libraries could be made or, if made, whether his single-chain TCRs could tolerate fusion of the bacteriophage coat protein. Chung does not report or suggest that the GPI or CD3ζ membrane anchor could be substituted with a bacteriophage coat protein. Even if there was such a teaching, there is no disclosure in the

reference about whether a recombinant bacteriophage could tolerate Chung's scTCR as part of the phage coat.

B. Barbas discloses heterodimeric receptor libraries that use phagmids. In the Abstract, phage are taught to encapsulate a genome encoding first and second polypeptides of a receptor such as an antibody; in which the first and second polypeptides are integrated into the coat matrix of the phage. Barbas generally discloses that such phage may include a polypeptide with an "insert domain" that has a receptor domain flanked by a secretion signal domain and a phage coat protein membrane anchor domain. Col. 14, lines 10-14. Heterodimeric receptors are preferred. See Col. 3, lines 1-41; Col. 14, lines 15-29; and Col. 15, lines 28-32. According to Barbas, there was some uncertainty in the field about which portions of bacteriophage coat proteins were needed for phage assembly. Col. 2, lines 19-46.

Barbas does not teach how to make or use a scTCR with or without a fused bacteriophage coat protein.

C. Onda reports use of a phage display system to explore binding interactions between the V-α region and antigen. Onda did not disclose use of the system to study TCR or single-chain TCR interactions. In the Abstract, Onda provides at pg. 1387:

> We utilized an M13 phage display system, designed for multivalent receptor display, to explore specific binding interactions between various TCR α chains and specific antigen in the absence of MHC.

That is, Onda fused **only** the V-α region to bacteriophage coat protein. The constructs **do not** include a V-β region and are **not scTCR fusion proteins.** Onda's fusions are much smaller TCR "half-molecules" lacking the V-β region and antigen binding pocket of Appellants' scTCR.

Onda at pg. 1395, col. 1, cautioned that his constructs were **unusual** and not typical of standard TCR interactions:

> Our results extend these findings by demonstrating that the dominant interactions of certain TCRα chains for peptide antigens may be sufficiently high that they can be analysed independently. However, these interactions are quite unusual in that they do not require the expression of the second TCR subunit or normal MHC and coreceptor interactions. **These results may raise concern that this model does not reflect typical TCR-ligand interactions.**

Significantly, only some of Onda's Vα chain fusion proteins were reported to bind antigen when fused to bacteriophage coat protein. At pg. 1395, col. 2 he states that:

> ...**only a subset of TCR Vα have capacity for direct interactions with antigen strong enough to be detectable in this system.**

Onda does not teach or suggest fusing a scTCR (V-α and V-β) to a bacteriophage fusion protein.

D. Huse described a phage vector system for screening and producing antibody F(ab) fragments.[5] Huse's system was reportedly used to produce free F(ab) and F(ab) displayed on the surface of bacteriophage. According to Huse however, not all attempts to produce F(ab) were successful. In more than a few instances, the recombinant bacteriophage made to produce the fusions apparently would not tolerate certain amounts of antibody protein. In describing attempts to display certain antibody H and L chains with his phage vector system, Huse stated on pg. 3919, col. 2 that:

> Phage titers of [phage vector] infected cultures were found to **decrease** relative to the level of F(ab)-pVIII fusion protein incorporation (cite omitted). Taken together, these results suggest that a **functionally viable phage particle may be able to tolerate a limited number of incorporated F(ab)-pVIII fusion products and that the amount of F(ab) incorporated into the phage coat may adversely affect phage titers and overall F(ab) yield.**

Huse does not disclose fusing a bacteriophage coat protein to a scTCR or TCR.

II. Summary of the Examiner's Argument

Grounds for the present rejection under 35 U.S.C §103 were formulated in the Office Action dated August 25, 2000. See paragraph B, above. The basis for the rejection has not changed substantially in the face of Appellants' arguments and claim amendments.

According to Appellants' understanding of the Examiner's alleged prima facie case, the primary references, Chung and Barbas, are alleged to teach a scTCR linked to a bacteriophage fusion protein. Onda and Barbas are relied on to teach that TCR-bacteriophage fusion proteins can be used to study antigen binding. Huse is used to teach that fusion proteins with the coat

[5] F(ab) is an abbreviation for an antigen binding fragment of an antibody (fragment antigen binding). F(ab) is a heterodimer consisting of two different chains ie., the antibody light and heavy chain. F(ab) is readily made by cleaving whole antibodies with specific proteolytic enzymes.

protein can be made in bacteria. The foundation of the Examiner's position is that because Barbas, Onda, and Huse teach some bacteriophage coat protein fusions, then it would be obvious to make fusions with Chung's scTCRs. Although facially somewhat plausible, the rejection is flawed on both scientific and legal principles as discussed herein.

III. The Examiner Erred in Rejecting Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72 as Being Obvious

A. Requirements of the prima facie case and its maintenance.

The Examiner erred in maintaining the obviousness rejection in the face of claim amendments and the state of the art as submitted made in this case and its parent. The Federal Circuit has reiterated that an Examiner's prima facie case is but a procedural tool of patent examination, with the express purpose of allocating the burdens of going forward as between the Examiner and Applicant. See In re Deckler 977 F.2d at 1449, citations omitted):

> Specifically, when obviousness is at issue, the examiner has the burden of persuasion and therefore the initial burden of production. Satisfying the burden of persuasion, constitutes a so-called prima facie showing. Once that burden is met, the applicant has the burden of production to demonstrate that the examiner's preliminary determination is not correct. The examiner, and if later involved, the Board, retain the ultimate burden of persuasion on this issue.

Clearly, as demonstrated herein, adequate evidence of the unobviousness of the claimed invention was provided by Appellants to shift the burden of persuasion to the Examiner.

In view thereof, it is requested that the Board review the obviousness question based on the invention as claimed, and the cited references, including all relevant parts thereof.

B. Standard For Reviewing An Obviousness Rejection under 35 USC §103.

The Federal Circuit has reiterated the manner in which obviousness rejections are to be reviewed. Where claimed subject matter has been rejected as obvious in view of a combination of prior art references, "a proper analysis under section 103 requires, inter alia, consideration of two factors: (1) whether the prior art would have suggested to those of ordinary skill in the art that they should make the claimed composition or device, or carry out the claimed process; and (2) whether the prior art would also have revealed that in so making or carrying out, those of

ordinary skill would have a reasonable expectation of success." In re Vaeck, 947 F.2d 488, 493, 20 U.S.P.Q.2d 1438, 1442 (Fed. Cir. 1991), cited In re Dow Chemical Co., 837 F.2d 469, 473, 5 U.S.P.Q.2d 1529, 1531 (Fed. Cir. 1988). As the Federal Circuit emphasized by succinctly summarizing: "Both the suggestion and the reasonable expectation of success must be founded in the prior art, not the Applicants' disclosure." Id. See also In re Merck & Co., Inc., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986).

More recently, the Federal Circuit has reviewed the case law regarding 35 U.S.C. §103. See In re Sang-Su Lee 277 F.3d 1388, 61 U.S.P.Q.2d 1430 (Fed. Cir. 2002).

Should the Board adopt the Examiner's prima facie case, Appellants submit that the claimed invention would not have been obvious in view of the legal standard summarized above.

C. **No Prima Facie Case of Obviousness**

As noted above, the foundation of the Examiner' argument rests on the belief that it would be obvious to fuse Chung's scTCR to the bacteriophage coat protein of Barbas because, allegedly, Barbas and Onda teach TCR-bacteriophage coat fusion proteins and Huse discloses that fusion protein with such a coat can be made in bacteria.

For the Examiner's prima facie case to stand, it is imperative that he establish that: 1) The cited references disclose or suggest fusing a bacteriophage coat protein to Chung's scTCR; 2) there is a settled role for the bacteriophage coat protein in making fusion proteins; and that 3) one could fuse Chung's scTCR to Barbas' coat protein with a reasonable expectation of success.

The Examiner's position is not supported by any of these points. Barbas, as relied on, does not teach or suggest a scTCR or even fusion of a scTCR to a bacteriophage coat protein. Onda and Huse, when read in their entirety, exemplify uncertainty in the field about using the bacteriophage coat protein to make certain fusion proteins. Even Barbas admitted that there was some doubt about how much one could change certain bacteriophage coat proteins without hindering phage assembly. Moreover, some of Onda's and Huse' molecules did not work well. Others did not work at all. On top of that uncertainty is heaped additional doubt about whether

Chung's "anchor" fragments could be substituted with the bacteriophage coat protein of Barbas. There was also doubt about whether the bacteriophage would tolerate fusion of the scTCR to its coat.

i) *Barbas and Onda does not teach or suggest a single-chain TCR (scTCR)*

In the Office Action dated August 25, 2000, Examiner Schwadron took the position that Barbas discloses:

> soluble fusion protein comprising a bacteriophage coat protein fragment covalently linked to a **single-chain heterodimeric receptor** (see the abstract and column 15, lines 27-28, in particular). Barbas also discloses that the **fusion protein may comprise domains of heterodimeric proteins derived from several ligand binding proteins,** including immunoglobulins and T cell receptors (see column 17, lines 62-66 and column 19, lines, 9-28. Barbas discloses that T cell receptor comprises alpha and beta chains each having a variable(V) and constant(C) region and T cell receptor has similarities in genetic organization and function to immunoglobulins (see column 19, lines 19-22, in particular).
>
> * * * *
>
> **Thus Barbas discloses but does not exemplify a soluble fusion protein comprising a bacteriophage coat protein covalently linked to T cell receptor domains**

The heterodimeric receptor proteins pointed out by the Examiner are not scTCRs. Heterodimeric proteins, and particularly the TCR of Barbas, are understood in the field to consist of two different α and β chains. Brief at pg. 1. Unlike the TCR, the scTCR of Appellants' claimed invention is a single-chain molecule with a V-α chain fused to a V-β chain. The position that Barbas discloses a "single-chain heterodimeric" receptor simply makes no sense. How can a single-chain molecule be a "heterodimer" when that requires two (dimer) different (hetero) chains? Brief at pg. 1 and footnote 1. Barbas could not have had the single-chain constructs of Onda and Chung in mind. Those references were published well after the priority date of the Barbas patent. Thus, nowhere in the reference is there any disclosure about how to make or use a scTCR.

Faced with this rebuttal, the Examiner took the position in the Final Office that Barbas' disclosure of "polypeptides comprising an insert domain" and "receptor proteins" should be read to include Chung's scTCRs. Also included in that sweeping reading of Barbas are "single chain or heterodimeric or single chain heteromers". That position is without merit. Too much is read from Barbas. It does not provide for any scTCR molecules. If the Examiner's overly-broad view

of the patent is allowed to stand and sweep in scTCRs, even though Chung's were published well after Barbas' priority date, it will preempt any attempt to obtain patent protection for scTCR-bacteriophage coat fusion proteins. A principle focus of Barbas was to provide **heterodimeric** receptors linked to a phage coat protein. See the Title of the patent, the Abstract and col. 3, lines 1-41, for instance. Such receptors are not the fusion proteins Appellants claim and there is no suggestion in Barbas to make or use them.

Even assuming, *arguendo*, that the Examiner is correct and that Barbas taught or suggested a scTCR (years before Chung or Onda were published), one reading Barbas in that way would be confused in light of the accepted understanding in the field that a heterodimer such as the TCR is a complex of two different polypeptide chains. Brief at pp. 1 – 2 and footnote 1.

As captioned above, Onda does not disclose a TCR or scTCR fusion to bacteriophage coat protein as alleged by the Examiner in the August 25, 2000 and June 17, 2002 Office Actions. Instead, Onda reports fusion of **TCR α chains** to bacteriophage coat protein. The TCR α chain is merely a part of the larger scTCR Appellants worked with. That is, the prior constructs are significantly smaller (and less likely to cause solubility problems when fused to coat proteins) than the scTCR fusions Appellants successfully made.

Moreover, the Examiner ignored Onda's clear hesitation about reading too much from TCR α chain constructs that include a fused bacteriophage coat protein. According to Onda, the interactions of the constructs were **unusual and not typical of TCR-ligand interactions**. See above and Onda at pg. 1395, col. 1.

Importantly, only some of Onda's TCR α chain constructs even worked to bind antigen. See above and pg. 1395 of Onda at col. 2, second full paragraph.

According to Onda then, some TCR α chain-bacteriophage coat protein fusions work and some do not. Those that do work were viewed as "unusual" and "not typical". In view of this caution, one working in this field would not be encouraged to fuse a bacteriophage coat protein

to a scTCR. None of the other cited references shed any light on Onda's clear hesitation to extend there findings to other TCR molecules.

The Examiner thus erred in trying to formulate a prima facie case by not giving due weight to all relevant portions of Onda. Contrary to this practice, it is well established that the Examiner must consider all relevant portions of cited references, including those portions which substantially weaken her position. In particular, the former CCPA stated in In re Mercier 515 F.2d 1161, 185 USPQ at 778:

> The relevant portions of a reference include not only those teachings which would suggest particular aspects of an invention to one having ordinary skill in the art, but also those teachings which would lead such a person away from the claimed invention.

See also Phillips Petroleum Co. v. U.S. Steel Corp., 673 F.Supp. 1278, 1315, 6 USPQ2d 1065, 1093 (D.Del. 1987), *aff'd*, 865 F.2d 1247, 9 USPQ2d 1461 (Fed. Cir. 1989).

The Board is thus urged to take Onda in its entirety and to consider all relevant portions of it including the passages quoted above. Read in this way, as it should, the reference would lead one in this field to doubt whether it would be feasible to fuse a bacteriophage coat protein to a scTCR to produce a soluble and functional fusion protein.

ii) *Huse reported difficulties producing some bacteriophage coat protein fusions*

The Huse reference, as quoted above, reported that not all F(ab)-pVIII (bacteriophage) coat proteins could be made at high titre. That is, Huse stated that the bacteriophage may not tolerate some amounts of F(ab) constructs, thereby decreasing phage titres and overall F(ab) yield. See above and Huse at pg. 3919, col. 2. When Huse is read in its entirety, as it should, the Examiner's statement that "Huse et al. teach that fusion proteins comprising bacteriophage VIII coat protein can be produced in bacteria" is an unsupported generalization. Huse clearly found that some amounts of heterodimeric F(ab) constructs harmed the bacteriophage that carried them. In view of this warning, a worker in the field would have good reason to doubt whether a bacteriophage could be fused to a scTCR or even a heterodimer such as a TCR without considerable experimentation.

The Examiner took the position in the Final Office Action that Barbas and Chung provide a reasonable expectation that one could make the claimed fusion proteins:

> Regarding **reasonable expectation of success**, both Barbas et al. and Chung et al. disclose use of phage display systems to produce single chain antibodies (see column 2, third paragraph from bottom and page 12658, first column).

However as clearly illustrated by Huse, not all phage display systems using antibodies work as expected. Some amounts of heterodimeric F(ab) antibodies cause problems. Thus the Examiner's position is not supported by the art of record in this case.

The Board is thus requested to take Huse in its entirety and to consider all relevant portions of it including the passage quoted above. In re Mercier 515 F.2d 1161, 185 USPQ at 778; and Phillips Petroleum Co. v. U.S. Steel Corp., 673 F.Supp. 1278, 1315, 6 USPQ2d 1065, 1093 (D.Del. 1987), *aff'd*, 865 F.2d 1247, 9 USPQ2d 1461 (Fed. Cir. 1989).

The substantial uncertainties raised by Onda and Huse have not been addressed by the Examiner. No objective scientific work has been made of record to resolve or explain them. Read in their entirety, as they should, Onda and Huse point out problems about making and using some bacteriophage coat protein fusions. Even if one skilled in this field were to read Onda and Huse selectively and disregard their warnings, there is still nothing in the art relied on to suggest that one could make or use a scTCR fusion to a bacteriophage coat protein. Even Barbas admitted that there was uncertainty about what coat protein parts could be manipulated for phage assembly.

In marked contrast, Appellants have demonstrated that it is possible to fuse a bacteriophage coat protein to the scTCR and obtain fully soluble and functional fusion protein. See Appellants' patent specification at Example 1 (showing construction of soluble scTCR fusion proteins); Examples 2-3 (production of special vectors to make the scTCR fusion proteins); Example s 4-5 (expression of Appellants' soluble scTCR fusions); Example 6 (purification of the soluble scTCR fusion proteins); Examples 7-11 and 16 (characterization of particular scTCR fusion proteins); and Example 15 (analysis of a bacteriophage library expressing Appellants' scTCR).

Accordingly, the §103 rejection fails both prongs of the Federal Circuit test for determining obviousness. See In re Vaeck, *supra*; and In re Dow Chemical Co., *supra*. It is submitted that the Board reverse the obviousness rejection in light of this test.

Whether or not the Examiner is taking the position that it would be **obvious to try** to make the claimed scTCR fusion proteins, both the Board and Federal Circuit have made it quite clear that this is not a burden that Appellants must bear. In particular, the Court in In re O'Farrell, 7 USPQ 2d 1673 (1988) held at page 1681:

> The admonition that "obvious to try" is not the standard under §103 has been directed mainly at two kinds of error. In some cases, what would have been "obvious to try" would have been to vary all parameters or try each of numerous possible choices until one possibly arrived at a successful result, where the prior art gave either no indication of which parameters were critical or no direction as to which of many possible choices is likely to be successful.

See also Ex parte Old, 229 USPQ 196, 200 (1985).

While the Court in In re O'Farrell went on to state that while obviousness does not require absolute predictability of success, what is required under §103 is a reasonable expectation of success.

Thus whether those in the field may have been tempted to fuse a bacteriophage coat protein to a scTCR, the field would have been cautioned from doing so in view of the warnings of Onda, Huse and to some extent even Barbas. The Barbas patent does not disclose or suggest any scTCR fusion to the coat protein. Chung does nothing to remedy these defects. In short, the field's unsuccessful experience with some amounts of antibody heterodimers (Huse) and some single-chain constructs (Onda), provides at worst no basis for believing that fusion of a bacteriophage coat protein to an scTCR will work and at best, a reason to doubt that such a fusion will result in a fully soluble and functional protein.

iii) *No teaching or suggestion that Chung's "anchor" fragment could be substituted with Barbas' bacteriophage coat protein*

Chung reported TCRs linked to a cell membrane anchor (glycosysl phosphatidylinositol (GPI) or murine CD3 ζ chain). The anchor apparently helps to express the single-chain TCRs. The anchor molecules are entirely different from the coat proteins of Barbas both in terms of chemical structure and function. For example, Chung's anchors are hydrophobic cell membrane proteins while those of Barbas are relatively more hydrophilic bacteriophage coat components. Chung's anchor apparently plants the scTCR in the membrane while the coat envelops the phage.

The Examiner has pointed to no teaching or suggestion in the cited art that Chung's anchor molecules could be substituted with Barbas' bacteriophage coat proteins. The obviousness rejection falls far short of establishing any nexus between Chung's anchors, which are attached to his scTCRs, and the coat proteins reported by Barbas.

iv) *Objective Evidence of Non-obviousness*

In addition to the lack of a prima facie case of obviousness, the strong objective evidence of non-obviousness presented during prosecution of this case further compels allowance of the claims.

Evidence of such objective indicia of non-obviousness, the so-called "secondary considerations" must be considered in all obviousness determinations. Stratoflex, Inc. v. Aeroquip Corp., 713 F.2d 1530, 1538-1539 (1983):

> Indeed, evidence of secondary consideration may often be the most probative and cogent evidence in the record. It may often establish that an invention appearing to have been obvious in light of the prior art was not. It is to be considered as part of all the evidence, not just when the decision-maker remains in doubt after reviewing the art.

See also Graham v. John Deere, 383 U.S. 1, 148 USPQ 459 (1966).

This standard set forth by the Federal Circuit applies not only during litigation of issued patents, but to a determination of patentability during ex parte prosecution as well. In re Sernaker, 702 F.2d. 989 217 USPQ 1, 7 (Fed. Cir. 1983). However, in the instant case, the Examiner is not properly considered evidence of "long felt need and failure of others" in maintaining the present § 103 rejection.

Specifically, Appellants' provided the Holler reference as indicating that the field longed to make the claimed fusion proteins but could not. Holler reported that phage display had not yet proven successful in making scTCRs despite what he saw as extensive structural similarity between antibodies and TCR V regions. Appellants' invention addressed this need and succeeded by providing soluble fusion molecules with a bacteriophage coat protein linked to the scTCR. The Holler reference is highly probative of the difficulties the field had in making these molecules and should be given substantial weight by Examiner Schwadron. MPEP 716.01(b).

The Examiner completely dismissed the Holler reference on grounds that "**Holler et al. simply not familiar with the prior art**". See Part D, above. That is no basis for disregarding the Holler's statement that the field wanted but failed to produce the claimed invention. Appellants are under no burden to provide evidence of Holler's knowledge of the art in order to have the reference considered as objective indicia of non-obviousness. See Stratoflex, Inc. v. Aeroquip Corp., 713 F.2d 1530, 1538-1539 (1983); In re Sernaker, 702 F.2d. 989 217 USPQ 1, 7 (Fed. Cir. 1983); and MPEP 716.01(b).

Moreover, Examiner Schwadron's citation of Weidanz et al. to support his disregard of Holler is clearly improper. That reference is not prior art and cannot serve as a basis for ignoring Holler or substantiating the obviousness rejection on appeal.

It is requested that the Board consider Holler as objective evidence that workers in the field wanted, but could not make, the claimed invention.

CONCLUSIONS

For the Examiner's prima facie case to stand, he has the burden of showing that: 1) The cited references disclose or suggest fusing a bacteriophage coat protein to a scTCR; 2) there is a settled role for the bacteriophage coat protein in making fusion proteins; and that 3) one could fuse Chung's scTCR to Barbas' coat protein with a reasonable expectation of success. These points have not been made by the Examiner. As discussed above, Barbas does not disclose scTCRs. Moreover, there was significant uncertainty in the field about whether it was possible to fuse a bacteriophage coat protein to a scTCR as exemplified by Huse and Onda. In view of the cited art and in consideration of the Examiner's position, it could be argued that one might be motivated to test fusing the bacteriophage coat protein to Chung's scTCR in the hope of producing a soluble and function protein. But this is not the legal standard required by our case law. It is without a doubt not obvious from the art of record to make the claimed invention of a scTCR fused to a bacteriophage coat protein.

Appellants submit that they have overcome the Examiner's obviousness rejection in the view of all the facts and argument of record in this case. Simply put, one of skill in this area would not be able to predict, with any reasonable expectation of success, how to make and use the claimed invention.

Importantly, Appellants have provided experimental evidence clearly showing that it is indeed possible to make and use scTCR-bacteriophage coat protein fusions. See Examples 1-11, 15 and 16 as discussed above.

In summary, Appellants submit that the instant invention is both novel and unobvious. The arguments set forth above establish that non-obviousness.

Although it is not believed that the present submission requires any fee for consideration by the Office, the Examiner is authorized to charge such fee to our deposit account **04-1105** should such fee be deemed necessary.

Respectfully submitted,

Date: August 5, 2004

By: _____

Robert L. Buchanan (Reg. No. 40,927)
EDWARDS & ANGELL, LLP
P. O. Box 55874
Boston, MA 02205
Tel. (617) 439-4444
Fax (617) 439-4170 / 7748
Customer No.: 21874

Doc. 453669

EXHIBIT A
Claims 1, 2, 4, 7, 8, 14, 67, 69, 71 and 72
on appeal

What is claimed is:

1. A soluble fusion protein comprising a bacteriophage coat protein covalently linked to a single-chain T cell receptor comprising an antigen binding pocket, wherein the single-chain T cell receptor comprises a V-α region covalently linked to a V-β region by a peptide linker sequence that effectively positions the V-α region and the V-β region to form the antigen binding pocket, the soluble fusion protein further comprising a C-β region fragment.

2. The soluble fusion protein of claim 1, wherein the C-terminus of the V-α region is covalently linked by the peptide linker sequence to the N-terminus of V-β region.

4. The soluble fusion protein of claim 2 wherein the C-β region fragment is covalently linked between the C-terminus of the V-β region and the N-terminus of the bacteriophage coat protein.

7. The soluble fusion protein of claim 2, wherein the peptide linker sequence contains from approximately 2 to 20 amino acids.

8. The soluble fusion protein of claim 1, wherein the bacteriophage coat protein is gene III or gene VIII protein.

14. A soluble fusion protein comprising covalently linked in sequence: 1) a V-α region, 2) a peptide linker sequence, 3) a V-β region covalently linked to a C-β region fragment, and 4) a bacteriophage gene VIII protein, wherein the peptide linker sequence effectively positions the V-α region and the V-β region to form an antigen binding pocket.

67. The soluble fusion protein of claim 1, wherein the C-terminus of the V-β region is covalently linked to the N-terminus of a C-β region fragment.

69. The soluble fusion protein of claim 1, wherein the V-α region and the V-β region are about 200 to 400 amino acids in length.

71. The soluble fusion protein of claim 1, wherein the C-β region fragment is about 50 to 126 amino acids in length.

72. The soluble fusion protein of claim 70, wherein the C-β region fragment does not include a cysteine residue corresponding to position 127 of a full-length C-β region.

#453669

0022-1767/92/14912-3914$02.00/0
THE JOURNAL OF IMMUNOLOGY

Vol. 149, 3914–3920, No. 12, December 15, 1992
Printed in U.S.A.

APPLICATION OF A FILAMENTOUS PHAGE pVIII FUSION PROTEIN SYSTEM SUITABLE FOR EFFICIENT PRODUCTION, SCREENING, AND MUTAGENESIS OF F(ab) ANTIBODY FRAGMENTS

WILLIAM D. HUSE,[1]* TIMOTHY J. STINCHCOMBE,* SCOTT M. GLASER,* LISA STARR,† MICHAEL MacLEAN,† KARL E. HELLSTRÖM,† INGEGERD HELLSTRÖM,† AND DALE E. YELTON*

*From *Ixsys, Inc., San Diego, CA 92121; †Department of Immunology, The Scripps Research Institute, La Jolla, CA 92037; and ‡Bristol-Myers Squibb Pharmaceutical Research Institute, Seattle, WA 98121*

We describe the application of a novel filamentous phage vector system suitable for efficient screening and production of F(ab) antibody fragments. The vector system can concurrently produce free F(ab) fragments and F(ab) displayed on the surface of M13 bacteriophage via a V_HC_H1-pVIII fusion protein. When expressed in a *supO* (nonsuppressor) strain of *Escherichia coli* free F(ab) can be produced. Antibody F(ab) fragments are secreted into culture medium at concentrations up to 0.3 mg/liter and conveniently subjected to detailed analysis with little or no purification. Higher concentrations of F(ab) (approximately 10 mg/liter) were found to accumulate in the periplasmic space. In this report the vector system is shown to produce correctly folded and assembled F(ab) fragments of chimeric L6, a mAb against a tumor-associated Ag expressed by many human carcinomas.

Until recently mAb have been primarily produced in mammalian cells. The slow growth rates and difficulty in genetically manipulating antibody genes expressed in mammalian cells have motivated development of methods to express antibody genes in simpler organisms. Molecular cloning techniques in bacteria have facilitated the production and manipulation of antibody fragments, increasingly aiding the search for useful antibodies (1, 2). Methods have been reported for the expression of antibody fragments in *Escherichia coli* using plasmids (3–7), bacteriophage λ (8–11), and more recently the filamentous phage M13 (12–17). Very large combinatorial libraries of 10^5 to 10^6 distinct antibody specificities can be created in microorganisms, far greater than can be achieved with hybridoma cell fusion methods (9, 10). However, the rapid identification, isolation, and, if necessary, modification of antibodies with the goal of improving affinity or redirecting specificity for Ag depends heavily on the availability of powerful screening methods both in terms of sampling large numbers of antibody fragments and evaluating certain aspects of antibody binding. Hence, an antibody expression vector that permits rapid cloning and mutagenesis of antibody V region genes and produces sufficient levels of antibody for biochemical analysis would be highly desirable.

Because antibody fragments can be displayed on the surface of filamentous phage, bacteriophage M13 vectors are proving particularly valuable in creating and screening sequence libraries for antibody fragments of interest. Briefly, fusion proteins are created by inserting DNA encoding an antibody fragment in front of a phage coat protein gene (18, 19). The fusion proteins become anchored in the phage coat via the coat protein and antibody sequence is displayed at the phage surface. Phage-bearing antibody sequences of interest can be detected by Ag binding and isolated in infectious form (13, 14). In addition to antibodies the incorporation of malarial protein (20), growth hormone (21), and a hexapeptide library (22) into the coat proteins of filamentous phage has been reported.

We have developed an M13 filamentous phage vector system that can produce and display F(ab) as a fusion product to pVIII coat protein and can also synthesize free F(ab) in quantity. We report here the production of chimeric L6 antibody F(ab) fragments in these M13-derived vectors. L6 is a mAb against a tumor-associated cell surface Ag expressed by many human carcinomas (23). L6 has been shown capable of lysing cancer cells in vitro (24) and is currently the subject of clinical trials (25).

MATERIALS AND METHODS

Construction of bacteriophage M13IXL604 for expression of L6 F(ab). Restriction enzymes, calf intestine alkaline phosphatase, T4 polynucleotide kinase, T4 DNA polymerase, and T4 DNA ligase were purchased from Boehringer-Mannheim (Indianapolis, IN). Total RNA was isolated from the chimeric L6-secreting cell line described by Feil et al. (26) by the guanidinium thiocyanate-phenol/chloroform method (27). First strand cDNA was synthesized using oligo dT and BRL Superscript reverse transcriptase (GIBCO BRL, Grand Island, NY) and PCR[2] amplification of L6 H chain (V_H-C_H1) and L chain (V_L-C_L) sequences was performed by the method of Saiki et al. (28) as modified by Sastry et al. (29). The following primers were used for PCR amplification. The restriction endonuclease recognizing the boldface and underlined cloning site in the sequences is indicated within the parentheses: forward V_H primer (*Xho*I): 5′-CAGTCTGGA-CCTGAG**CTCGAG**AAGCCTGGAGAG-3′; forward V_L primer (*Nco*I)

Received for publication June 17, 1992.
Accepted for publication September 25, 1992.
The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *divertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Address correspondence and reprint requests to Dr. William D. Huse, IXSYS, Inc., Suite L103, Bldg. 2, 3550 General Atomics Court, San Diego, CA 92121.

[2] Abbreviations used in this paper: PCR, polymerase chain reaction; gVIII, gene VIII coding for the major coat protein of M13 phage; pVIII, major coat protein of M13 phage; C_H1 H chain C region 1.; V_L, L chain V region; IPTG, isopropyl-β-D-thiogalactopyranoside; V_H, H chain V region; MOPS, 3-(*N*-morpholino)propane sulfonic acid; TES, 30 mM Tris-HCl, 2 mM EDTA, 2% sucrose (w/v), pH 8.0; *supE*, strains of *E. coli* that carry a glutamine inserting amber (UAG) suppressor tRNA; MES, 20 mM 3-(*N*-morpholino)propane sulfonic acid, 2 mM EDTA, 20% sucrose (w/v), pH 7.5; *supO*, strains of *E. coli* that do not carry a suppressor tRNA.

5′-GCCCAACCAGCCATGGCCCAAATTGTTCTCTCCCAGTCT-3′; reverse C_H1 primer (*Spe*I): 5′-TGTGTGAGTACTAGTACAAGATTT-GGGCTCTGC-3′; reverse C_κ primer (*Xba*I): 5′-CCGCTTAACTCT-AGACTAACACTCTCCCCTGTTCAAGCT-3′. The V_H-C_H1 and V_L-C_κ chains were digested and ligated into prepared M13IX31[3] and M13IX12 vectors (30) resulting in M13IX31L6-H and M13IX12L6-L, respectively (Fig. 1). Sequence corrections of the cloned L6 V_H-C_H1 and V_L-C_κ chains were accomplished by site-directed mutagenesis of uracil-substituted ssDNA as described (31, 32). The correcting nucleotides are indicated by the underlined sequences. Five N-terminal V_H amino acids were inadvertently omitted from the original L6 sequence information resulting in an 11-amino acid deletion upon cloning. One V_H correction primer was 5′-CTCTCCAGGCTTCTTC-AGCTCAGGTCCAGAGGCTTTTG(C/T)CAC-3′. This primer corrected the *Xho*I cloning site back to the original L6 sequence and replaced V_H codons 7 to 10. The C/T mixed site introduced an amino acid change in the leader sequence that was found to increase expression. The following V_H primer served to restore the remaining six N-terminal amino acids 5′-CAGCTCAGGTCCAGACTGCACCAAC-TGGATCTGGGCCATGGCTGGTTGGGC-3′. The following V_L primer served to correct an incorrect nucleotide in the reverse C_κ PCR primer: 5′-ACTCTCCCCTGTTGAAGCTCTTTGTGA-3′. The H chain encoding M13IX31L6-H and L chain encoding M13IX12L6-L vectors were combined by annealing as described (30) to form M13IXL604 (Fig. 1). M13IXL605 was derived from M13IXL604 by mutation of the TAG (stop) codon, located between the H chain encoding sequence and the pseudo wild-type gene VIII, to GGT (glycine) using the oligonucleotide 5′-CGCCTTCAGCACCGGATCCACTAGT-3′ so that continual V_H-C_H1-pVIII fusion protein would be made. DNA sequence analysis of ssDNA prepared from phage isolates was performed with Sequenase Version 2 according to the manufacturer (United States Biochemical, Cleveland, OH).

Antibodies and reagents. The anti-Id mAb to L6 have been previously described (33). Anti-Id 1 is a γ2b isotype, anti-Id 3 is a γ2a isotype, and both anti-Id 7 and anti-Id 13 are γ1 isotypes. Alkaline phosphatase-conjugated antibodies were purchased from Fisher Biotech (San Francisco, CA). Unconjugated goat anti-human κ antibody was purchased from Caltag Laboratories (So. San Francisco, CA). Rabbit anti-M13 IgG was purified by Sepharose-protein A chromatography and exhaustively absorbed against whole *E. coli*. The anti-M13 antibody was subsequently biotinylated using D-biotinoyl-ε-aminocaproic acid *N*-hydroxysuccinimide ester (Boehringer-Mannheim) using standard chemistries. Vectastain avidin-horseradish peroxidase complex (Vector Laboratories, Burlingame, CA) and streptavidin-alkaline phosphatase complex (Boehringer-Mannheim) were used for second step reactions.

Screening by replicate filter lifts. M13IXL604 phage were plated at a low plaque density. F(ab) expression was induced by overlaying the plate with a 0.45-μ nitrocellulose filter (Schleicher and Schuell, Keene, NH) soaked in 10 mM IPTG and incubating at room temperature from 6 h to overnight. The filter was removed and placed in blocking buffer (Biosite Diagnostics, San Diego, CA) to block nonspecific binding sites. Phage growth was resumed by incubating the plate for an additional 2 h at 37°C and a second filter applied as described above. This procedure was repeated for the last filter and all filters were then blocked in blocking buffer. All monoclonal or polyclonal antibodies to be used for screening were diluted in blocking buffer. Filters were probed with either alkaline phosphatase-conjugated goat antibody to human λ-chain, alkaline phosphatase-conjugated goat antibody to human κ-chain, or anti-Id 3, which binds to L6 antibody. In the case of anti-Id 3, a secondary alkaline phosphatase-conjugated goat antibody to mouse IgG2a was used for detection. All filters were then washed three times for 10 min with 25 mM Tris, 0.137 M NaCl, 5 mM KCl, 0.9 mM $CaCl_2$, 0.5 mM $MgCl_2$, and 0.05% Tween 20 (pH 7.4) and developed with alkaline phosphatase substrate reagent (Bio-Rad, Richmond, CA).

F(ab) production and purification. For analytic scale production of F(ab) the *supE* amber suppressor strain XL-1 (Stratagene, San Diego, CA) and the *supO* nonsuppressor strain MK30–3 (Boehringer-Mannheim) were each grown in 2X YT medium at 37°C until the cultures reached a density of 0.4 to 0.6 at OD_{600}. Each strain was then diluted 1/10 into three culture tubes containing 3 ml 2X YT and infected with 3 μl of high titer (10^{11} plaque-forming units/ml) phage stock of M13IXL604, M13IXL605, or M13IX31/tube and incubated with shaking for 3 h at 37°C. Protein synthesis was induced by the addition of IPTG to a final concentration of 1 mM and shaking allowed to proceed for 10 to 14 h at ambient temperature. Periplasmic fractions were prepared essentially as described by Skerra and Plückthorn (5). The infected cultures were centrifuged for 10 min at 4°C and the cleared culture supernatant was reserved. The bacterial pellet was resuspended in 60 ml cold TES and digested on ice for 10 min after adding an equal volume of a cold freshly prepared solution containing 2 mg/ml lysozyme in TES. The periplasmic space containing free F(ab) was fractionated by centrifugation at 9000 rpm for 10 min at 4°C. The soluble periplasmic fraction was retained and diluted in blocking buffer to adjust the concentration of F(ab) to be equivalent to that found in the cleared culture supernatant.

For purification of L6 F(ab) a 1-liter culture of MK30–3 was grown and infected with M13IXL604 as described above. The cells were harvested by centrifugation at 5800 × *g* for 10 min at 4°C. The pellet was resuspended in 20 ml MES at ambient temperature. An equal volume of 2 mg/ml lysozyme in MES was added with mild vortexing and the suspension incubated at ambient temperature for 10 min. The soluble periplasmic fraction was isolated by centrifugation at 9700 × *g* for 10 min at 4°C. The periplasmic fraction was subjected to a second centrifugation at 12,000 × *g* for 30 min at 4°C. The cleared periplasmic fraction was loaded onto a macroprep 50 S support (Bio-Rad) washed extensively with 20 mM MOPS, pH 7.5 and eluted with 20 mM MOPS, 120 mM NaCl, pH 7.5. The partially purified L6 F(ab) was concentrated by centrifugation in a Centricon 30 device (Amicon, Beverly, MA) and size fractionated on a Biogel P-60 (Bio-Rad) column equilibrated and eluted with 20 mM MOPS, 120 mM NaCl, pH 7.5. Eluted L6 F(ab) fractions were assayed for binding to anti-Id 3 by ELISA. Fractions eluting at 8, 9, and 10 ml were then pooled and concentrated. Samples from each purification step were analyzed by SDS-PAGE on a 10% nonreducing gel followed by staining with Coomassie brilliant blue.

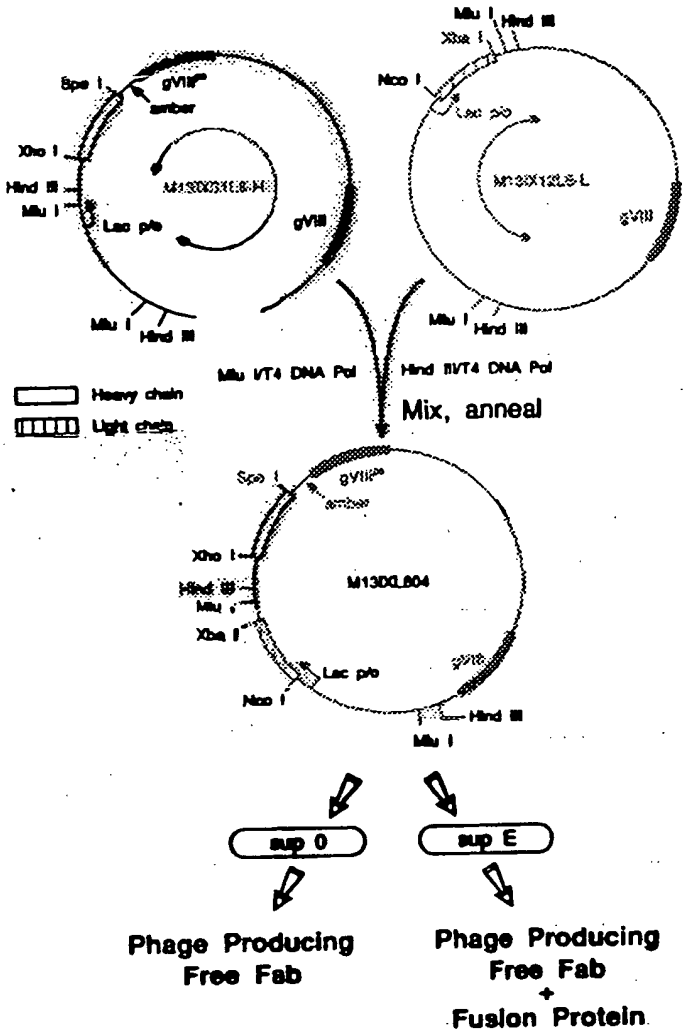


Figure 1. Construction of M13IXL604 vector for expression of L6 Fab. The M13DXL604 vector product contains one copy of pseudo wild-type gVIII (gVIII*) and one wild-type gVIII (gVIII) both downstream from the H chain encoding region. gVIII* has been altered in sequence to reduce homologous recombination with gVIII contained on the same vector. The presence of gVIII reduces the likelihood of selection against certain fusion proteins as the result of compromised phage viability (34).

[3] M13IX31 and M13IX12 vectors are available at no charge from Ixsys, San Diego, CA.

ELISA characterization of L6 F(ab) produced by M13IXL604. M13IXL604 phage stocks were prepared for assay by overnight infection of XL-1 in LB broth and 10 μg/ml tetracycline to maintain expression of F′, and 1 mM IPTG. The bacteria were removed by centrifugation (10 min in a microfuge) and the phage solution was assayed for binding to L6-specific anti-Id 1, 3, 7, 13, and to an anti-κ-chain antibody by ELISA. All antibodies were coated onto Immulon 1 microtiter plates (Dynatech Laboratories, Chantilly, VA) at 10 μg/ml in 0.1 M $NaHCO_3$, pH 8.5, overnight at 4°C. The antibody solution was shaken out and the plates were blocked with 300 μl of specimen blocking buffer (Genetic Systems, Seattle, WA) for 1 h at room temperature. Plates were washed before use with 0.5% Tween 20 in 0.15 M NaCl and with the same solution between incubations. Phage dilutions were dispensed into microtiter plates in 100 μl of blocking buffer and incubated overnight. After washing, 100 μl of conjugate blocking buffer (Genetic Systems) containing 1 μg/ml of biotinylated rabbit antibodies to M13 were incubated 1 h at ambient temperature. After washing, 100 μl of Vectastain avidin-HRP complex were incubated 30 min at room temperature. After a final wash step, 100 μl of chromogenic substrate (3,3′,5,5′-tetramethylbenzidine) in a citrate/phosphate buffer were added. The reaction was stopped with 100 μl of 3 N H_2SO_4 at various times ranging from 10 min to 2 h to achieve an optimum signal to background ratio for each different plate-coating antibody. Plates were read on a microplate reader (Biotek, Burlington, VT) in dual channel mode at 450/630 nm. All assay points were measured in duplicate. M13IX31 served as a negative control.

The expression of M13IXL604, M13IXL605, and M13IX31 in bacterial strains MK30–3 and XL-1 was also analyzed by ELISA. Culture supernatant and diluted periplasmic fractions prepared from infected XL-1 and MK30–3 bacteria described above were serially diluted in diluent (Biosite Diagnostics), added to microtiter plates coated with anti-Id 3 and incubated for 2 h at ambient temperature. L6 F(ab) expressed in the culture medium or in the periplasmic space was detected with an alkaline phosphatase-conjugated goat antibody to human κ-chains. Expression of L6 F(ab) displayed on the phage surface was detected by incubating the captured sample first with the biotinylated rabbit antibody to M13 and then with streptavidin-alkaline phosphatase complex. The washed plates were developed with 6 mg/ml phenolphthalein monophosphate in 0.1 M aminomethylpropanediol, 0.5 M Tris, and 0.1% NaN_3, pH 10.2 (JBL Scientific, San Luis Obispo, CA) for 10 to 30 min. The reaction was stopped by addition of one-third volume of cold 30 mM Tris base, 6 mM EDTA, and the absorbance at 560 nm measured.

RESULTS

Construction of L6 F(ab) expression vector M13IXL604. The construction of M13IXL604 is shown in Figure 1. Total RNA was isolated from the chimeric L6-secreting cell line and cDNA synthesized. After PCR amplification of the cDNA using sequence-specific primers, the chimeric L6 V_L-C_κ L chain was digested to completion with *NcoI* and *XbaI* and cloned into the *NcoI*/*XbaI* site of M13IX12 to construct M13IX12L6-L. Similarly, the chimeric L6 V_H-C_H1 H chain PCR product was digested to completion with *XhoI* and *SpeI* and cloned into the *XhoI*/*SpeI* site of M13IX31 to construct M13IX31L6-H. M13IX31 contains an amber stop codon located directly 5′ of a modified gVIII gene (pseudo-gVIII) encoding mature M13 pVIII major coat protein. Cloning antibody V_H or V_H-C_H1 regions into the *XhoI*/*SpeI* site of M13IX31 abuts these antibody sequences in frame with the amber stop-pseudo gVIII sequence. Thus, M13IX31L6-H should produce L6 V_H-C_H1-pVIII fusion product when expressed in an *E. coli* amber suppressor strain such as XL-1 and produce predominantly free L6 V_H-C_H1 protein when expressed in a nonsuppressor strain such as MK30–3. The H and L chain vectors were then recombined into the single M13IXL604 expression phage by annealing *HindIII*/T4 DNA polymerase-digested M13IX12L6-L to *MluI*/T4 DNA polymerase-digested M13IX31L6-H through the homologous regions found between the *HindIII* and *MluI* restriction sites contained in both vectors (30). Correct L6 H and L chain sequences were confirmed by DNA sequence analysis. The vector M13IXL605, which contains a GGT codon coding for glycine in place of the amber stop found in M13IXL604, was constructed by site-directed mutagenesis of M13IXL604. Thus, M13IXL605 should produce L6 V_H-C_H1-pVIII fusion product irrespective of the bacterial host used for expression.

Functional characterization and purification of M13IXL604-expressed L6 F(ab). M13IXL604 phage expressing L6 F(ab) was initially characterized by replicate filter lift assays (Fig. 2). XL-1 bacteria were infected at a low multiplicity of infection resulting in low plaque density of phage. No positive plaques were detected by filter lift assay when probed with alkaline phosphatase-conjugated goat antibody to human λL chain (Fig. 2*A*). In contrast, numerous, superimposable, uniform positive plaques were detected in assays of human κ-chain (Fig. 2*B*) and anti-Id 3, which recognizes assembled L6 H and L chain (Fig. 2*C*). Thus, screening by replicate filter lifts allows for detection of functional F(ab) expressed in this system.

To further evaluate the functional integrity of L6 F(ab) expressed by M13IXL604, phage displaying L6 F(ab) on the phage surface were assayed for binding to a panel of distinct mouse monoclonal anti-Id raised against the mouse mAb L6. ELISA experiments shown in Figure 3 demonstrated that four anti-Id antibodies specific for L6

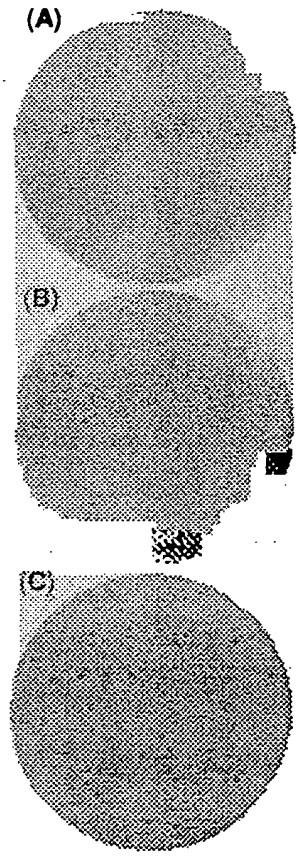


Figure 2. Screening for L6 F(ab) reactivity by replicate filter lifts. M13IXL604 phage was replated at low density for screening. Replicate filter lifts were prepared and probed with *A*) alkaline phosphatase-conjugated goat antibody to human λ-chain, *B*) alkaline phosphatase-conjugated goat antibody to human κ-chain, and *C*) anti-Id 3 antibody. The filter in *C* was washed and incubated with alkaline phosphatase-conjugated goat antibody to mouse IgG2a. Filters were then washed and developed with alkaline phosphatase substrate reagent.

are recognized by the fusion protein displayed on the M13IXL604 phage. In the ELISA format used, M13IXL604 phage were grown in the suppressor bacterial strain XL-1, and phage displaying incorporated L6 F(ab)-pVIII protein are captured by solid phase anti-Id antibodies and detected with anti-M13 antibodies. Thus, only phage-bound L6 F(ab) is detected. The four anti-Id antibodies exhibit different binding specificities to L6 (33). Anti-Id 1 and 7 bind to L6 L chain, 13 to H chain and 3 to assembled L and H chains. The binding of M13IXL604 phage to an anti-κ-chain antibody further confirmed the assembly of L6 L chain with L6 H chain-pVIII fusion protein to yield F(ab), because L6 L chains are of the κ type (Fig. 3). M13IX31, which contains no H chain-encoding region served as a negative control for each anti-Id. A representative curve shows no M13IX31 binding. The ELISA results in Figure 3 thus indicate the presence of both L6 L and H chains and confirm that proper folding and assembly of L and H chains on the phage surface have been achieved.

Purification of L6 F(ab) fragments produced by M13IXL604 phage grown in the nonsuppressor bacterial strain MK30–3 was accomplished by loading the crude periplasmic fraction prepared from a 1-liter shake flask culture onto a cation exchange resin. The eluted material was concentrated and size fractionated by gel exclusion chromatography. Figure 4*A* shows the elution profile from the gel exclusion column as assayed by ELISA for functional binding to anti-Id 3 using anti-human κ to detect captured L6 F(ab). The purified F(ab) was shown to have a M_r of 42.7 kDa by SDS-PAGE, a value consistent with the 48-kDa size of known F(ab) fragments (Fig. 4*B*). Based therefore on both functional and physical properties, the identity of the purified material is chimeric L6 F(ab).

Regulated expression of M13IXL604 production of F(ab) and phage displayed F(ab). The proper activity of M13IXL604-expressed L6 F(ab) having been established, the functioning of the expression control system of M13IXL604 was tested. This involved comparing levels of fusion F(ab) with free F(ab) produced when phage were grown in suppressor strains (*supE*) vs nonsuppressor

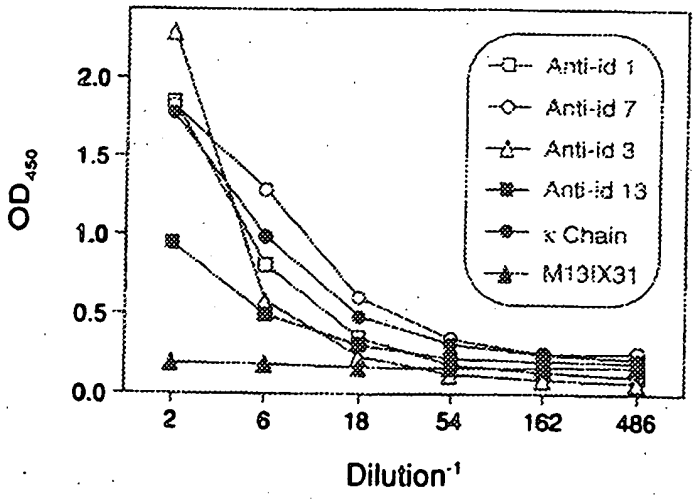


Figure 3. ELISA characterization of L6 F(ab) produced by M13IXL604. Various dilutions of phage expressing L6 F(ab) on the phage surface were evaluated for binding to anti-Id 1, 3, 7, and 13 and to an anti-human κ-chain antibody. Bound phage were then detected as described in *Materials and Methods*. Curves represent the average of triplicate assays.

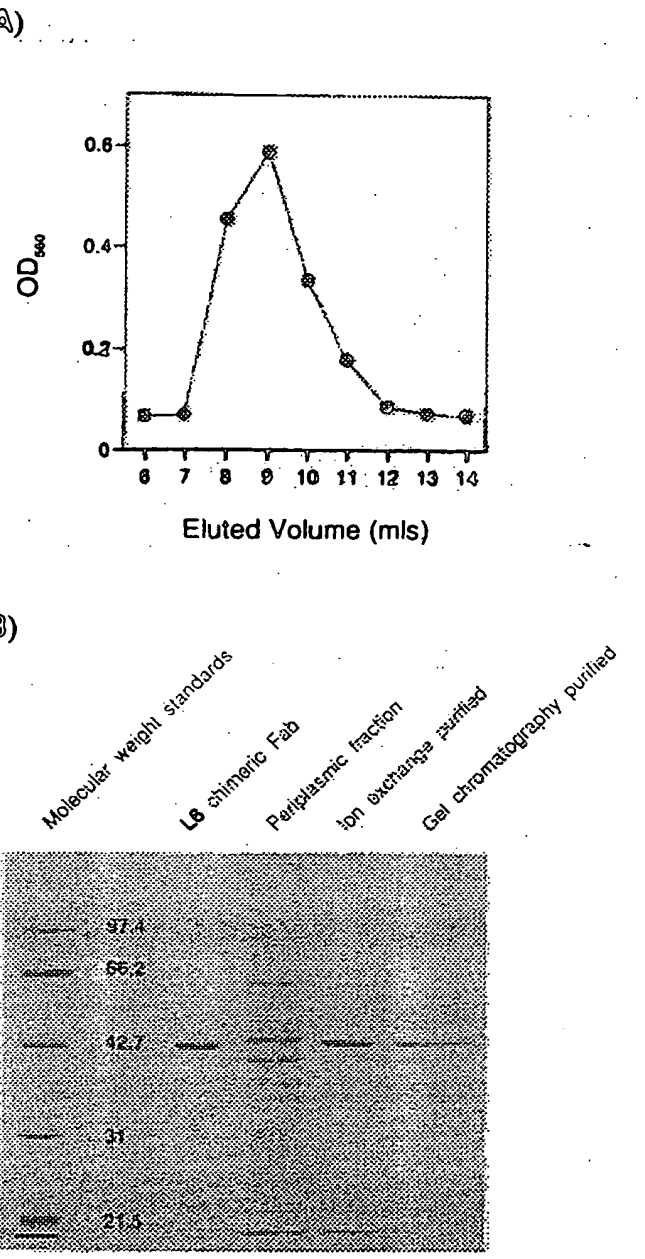


Figure 4. Purification of L6 F(ab) expressed by M13IXL604-infected MK30–3 bacteria. *A*) elution profile of L6 F(ab) fractions assayed for binding to anti-Id 3 antibody by ELISA. Fractions eluting at 8, 9, and 10 ml were then pooled and concentrated. *B*) SDS-PAGE analysis of purified Fab. Samples from each purification step were analyzed by SDS-PAGE on a 10% nonreducing gel followed by staining with Coomassie brilliant blue. The L6 chimeric F(ab) standard was prepared from whole IgG by papain digestion followed by protein A chromatography (Pierce, Rockford, IL).

strains (*supO*) of bacteria, respectively. This comparison was made using ELISA assays employing anti-Id 3 to detect assembled L6 F(ab). F(ab) displayed on the surface of M13 phage was detected using the M13 assay described in Figure 2. M13IXL605, a variant construct of M13IXL604 that has the amber stop codon replaced with a glycine encoding codon and therefore always synthesizes V_H-C_H1-pVIII protein, served as a control.

When phage were grown in the *supE* strain XL-1, both M13IXL604 and M13IXL605 secreted phage displaying F(ab) into the culture medium (Fig. 5*A*) and, as expected, little or no mature phage displaying F(ab) was detected in the periplasmic space (Fig. 5*B*). In the *supO* strain

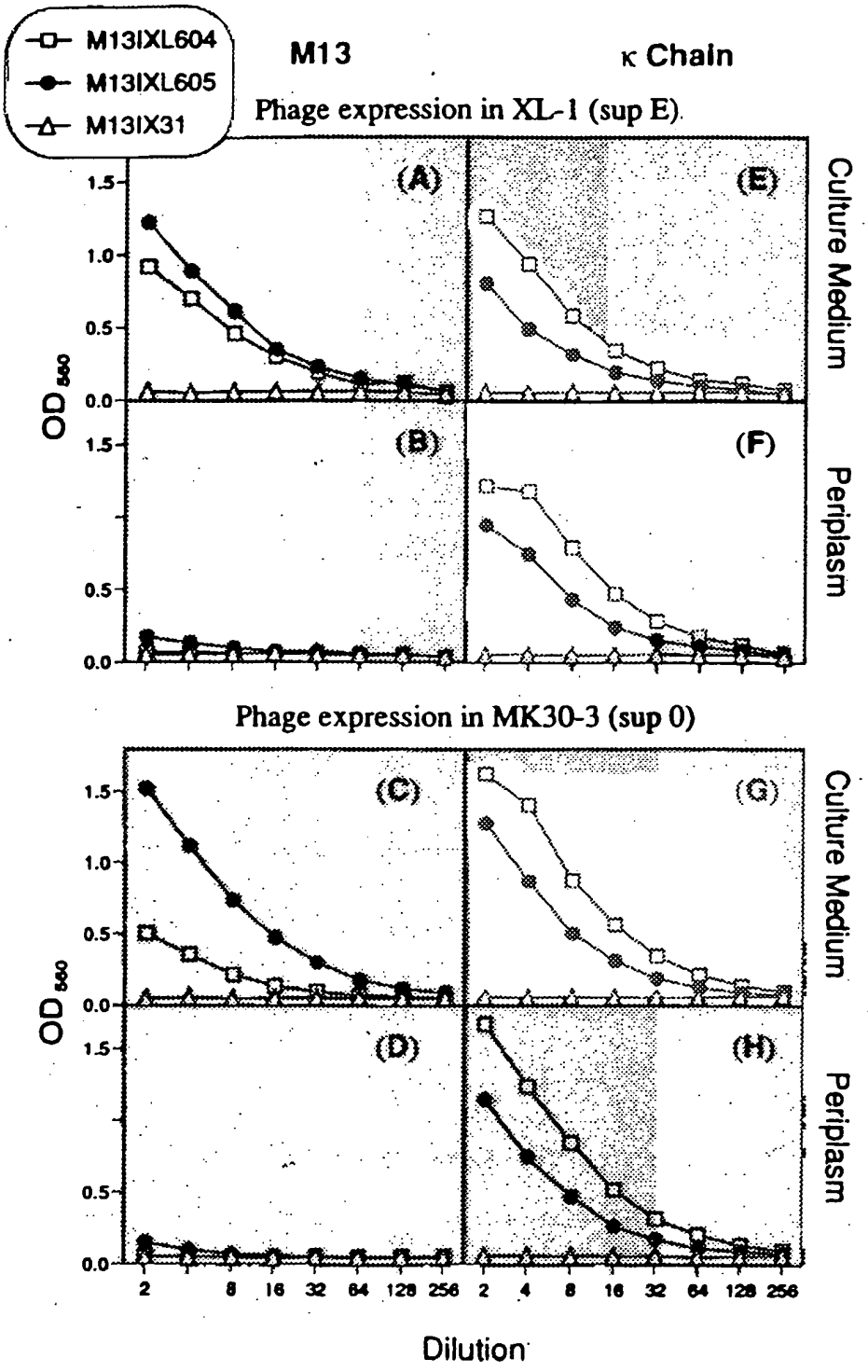


Figure 5. Characterization of the M13IXL604 expression control system. uid cultures of the amber suppres- strain XL-1 (*supE*) and the non- pressor strain MK30–3 (*supO*) were h infected with either M13IXL604, 3IXL605, or M13IX31 and culture su- natant and periplasmic fractions pre- ed as described in *Materials and hods*. Culture supernatant and peri- um from infected cultures were as- ed for functional binding to anti-Id 3 F(ab) displayed on phage surface (*A* to und by total F(ab) (*E* to *H*).

30–3, the amber stop codon in M13IXL604 should vent fusion protein formation. Figure 5*C* shows phage playing F(ab) to be sharply reduced in M13IXL604 vs 3IXL605. The low signal seen in the M13IXL604 cul- e medium is presumably caused by nonspecific asso- tion of L6 F(ab) with phage particles and/or transla- ıal readthrough. Again no mature phage displaying b) was found in the periplasm (Fig. 5*D*). In examining production of F(ab) by the anti-Id 3/κ-chain assay h M13IXL604 and M13IXL605 secreted anti-κ-reac- : material into the culture medium when expressed in ıer XL-1 or MK30–3 (Figs. 5*E* and 5*G*). Because the efficiency of amber suppression can be widely variable, we expect that M13IXL604 expressed in XL-1 should secrete free F(ab) in addition to F(ab) displayed on the surface of phage. Thus, the signal in Figure 5*E* should represent some combination of the two molecules. Polyethylene glycol precipitation of both M13IXL604 and M13IXL605 phage from induced XL-1 culture supernatants showed that approximately 95% of the anti-κ-reactive material remained in the polyethylene glycol-cleared supernatant (data not shown) indicating that the predominant secreted product is free L6 F(ab). Consistent with the view that M13IXL604 and M13IXL605 secrete free

L6 F(ab), abundant amounts of soluble assembled F(ab) were detected in the periplasmic space in both strains of *E. coli* (Fig. 5, *F* and *H*). MK30–3 appeared to consistently produce higher concentrations of F(ab) than XL-1. Quantitative ELISA demonstrated that the concentration of free F(ab) produced by M13IXL604 grown in the *supO* strain MK30–3 approaches values up to 0.3 mg/liter in culture supernatant and concentrations approaching 10 mg/liter within the periplasmic space (data not shown).

DISCUSSION

M13IX31 (for H chains) and M13IX12 (for L chains) are the M13 vectors that can receive polyclonal sets of V regions for constructing combinatorial libraries or individual antibody genes for V region production and/or mutagenesis (30). The system is the first to allow differentially controllable production of free and fusion F(ab). Upon recombining the two vectors, all of the control elements required for regulated expression of phage displayed F(ab) or free F(ab) are contained in a single expression vector. For phage display F(ab) is synthesized as a fusion product to pVIII coat protein, the major structural protein of the filamentous M13 phage particle. In the M13 bacteriophage the major coat protein pVIII is expressed in several thousand copies per phage particle. Although this level of expression may be useful for producing F(ab) as F(ab)-pVIII fusion proteins, pVIII synthesized solely as a fusion product is not likely to form a properly assembled coat as a result of steric hindrance by the much larger H chain polypeptide. We reasoned that if wild-type pVIII, in addition to F(ab)-pVIII fusion product, is available for assembly then the formation of a mature infectious phage particle could occur. With this in mind, the H chain-gVIII expression vector M13IX31 was constructed to contain two copies of gVIII. The copy that anchors F(ab) to the surface of M13 has had gVIII codons extensively altered (pseudo gVIII) to prevent recombination with the wild-type gVIII.

To demonstrate the utility of these vectors we cloned the antibody V_L-Cκ- and V_H-C_H1-encoding sequences from the chimeric L6 transfectoma cell line into the M13IX12 and M13IX31 vectors, respectively. The two vectors were recombined to create M13IXL604. M13IXL604 expresses a dicistronic message encoding both H and L chain L6 sequences under transcriptional control of an inducible Lac promoter. An amber stop codon resides between the H chain-encoding region and the pseudo gVIII coat protein. When grown in an amber suppressor strain (*supE*) of *E. coli*, M13IXL604 was shown to produce L6 H chain-pVIII fusion protein resulting in phage-displayed F(ab) in addition to free F(ab). When free F(ab) was desired as the major end product, the phage was grown in a nonsuppressor strain (*supO*). The M13IXL605 vector, which solely produces F(ab)-pVIII fusion protein, was constructed to evaluate, by comparison, the M13IXL604 control system.

M13IXL605 was found to display slightly higher levels of phage-associated F(ab) than M13IXL604 when grown in the *supE* bacterial strain. This is likely attributed to a higher proportion of F(ab)-pVIII fusion protein relative to wild-type pVIII produced in the M13IXL605-infected strain. In both the *supE* and *supO* bacterial strains M13IXL604 secreted chimeric L6 F(ab) at levels somewhat higher than those of M13IXL605. Phage titers of M13IX31-, M13IXL604-, and M13IXL605-infected cultures were found to decrease relative to the level of F(ab)-pVIII fusion protein incorporation (D. Yelton, unpublished observations). Taken together, these results suggest that a functionally viable phage particle may be able to tolerate a limited number of incorporated F(ab)-pVIII fusion products and that the amount of F(ab) incorporated into the phage coat may inversely affect phage titers and overall F(ab) yield. Furthermore, the appearance of free F(ab) in the culture media suggests that H chain-pVIII fusion protein unincorporated into phage is properly assembled with L chain and secreted as functional F(ab).

The set of anti-Id antibodies raised against the murine mAb L6 demonstrated that faithful expression and function of bacterially produced chimeric L6 F(ab) occurs in the M13IX12 and M13IX31 vector system. L6 F(ab) activity is readily detectable by either the nitrocellulose filter lift or ELISA formats. The anti-Id antibodies used in these studies serve as a convenient and informative model for antibody-Ag binding, inasmuch as the tumor Ag bound by L6 in vivo is yet to be purified or fully characterized.

The M13IX31 and M13IX12 vector system described here can serve as a versatile, general purpose approach to F(ab) production and screening. F(ab) production in this system is sufficiently robust to permit multiple replicate filter lifts, a practical requirement for implementing various screening strategies. The vectors can be used to create combinatorial antibody libraries to identify novel antibodies. Screening of sufficiently large combinatorial antibody libraries could potentially allow useful antibody fragments of murine or human origin to be isolated without the necessity to perform standard immunization procedures (35). In addition, because oligonucleotide-directed mutagenesis is convenient and highly efficient in M13 these vectors are ideal for engineering antibodies with new properties. It is conceivable that the ability to create stable, high affinity human antibodies is a promising endeavor.

In a companion publication (36), we show that the M13IX31 and M13IX12 antibody expression system lends itself to efficient antibody engineering by site-directed mutagenesis. Codon-based mutagenesis of L6 hypervariable regions proved effective in altering the fine specificity of L6 in a predefined manner. Thus the system allows mutagenesis, screening, and mutant F(ab) production to be accomplished very rapidly.

Acknowledgments. We thank G. Vaiklrs and members of Biosite Diagnostics for useful discussions and for supplying the formulation for reagents used in this study. We thank S. L. Morrison for review of the manuscript and G. A. Cruz for technical assistance.

REFERENCES

1. Plückthorn, A. 1990. Antibodies from *Escherichia coli*. *Nature 347:497.*
2. Winter, G., and C. Milstein. 1991. Man-made antibodies. *Nature 349:293.*
3. Boss, M. A., J. H. Kenten, C. R. Wood, and J. S. Emtage. 1984. Assembly for functional antibodies from immunoglobulin heavy and light chains synthesized in *E. coli*. *Nucleic Acids Res. 12:3791.*
4. Cabilly, S., A. D. Riggs, H. Pande, J. E. Shively, W. E. Holmes, M. Rey, L. J. Perry, R. Wetzel, and H. L. Heyneker. 1984. Generation of antibody activity from immunoglobulin chains produced in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA 81:3273.*
5. Skerra, A., and A. Plückthorn. 1988. Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science 240:1038.*

6. Better, M., C. P. Chang, R. R. Robinson, and A. H. Horwitz. 1988. Escherichia coli secretion of an active chimeric antibody fragment. *Science 240:1041.*
7. Ward, E. S., D. Güssow, A. D. Griffiths, P. T. Jones, and G. Winter. 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli. Nature 341:544.*
8. Sastry, L., M. Alting-Mees, W. D. Huse, J. M. Short, J. A. Sorge, B. N. Hay, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1989. Cloning of the immunological repertoire in *Escherichia coli* for generation of monoclonal catalytic antibodies: construction of a heavy chain variable region-specific cDNA library. *Proc. Natl. Acad. Sci. USA 86:5728.*
9. Huse, W. D., L. Sastry, S. A. Iverson, A. S. Kang, M. Alting-Mees, D. R. Burton, S. J. Benkovic, and R. A. Lerner. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science 246:1275.*
10. Caton, A. J., and H. Koprowski. 1990. Influenza virus hemagglutinin-specific antibodies isolated from a combinatorial expression library are closely related to the immune response of the donor. *Proc. Natl. Acad. Sci. USA 87:6450.*
11. Persson, M. A. A., R. H. Caothien, and D. R. Burton. 1991. Generation of diverse high-affinity human monoclonal antibodies by repertoire cloning. *Proc. Natl. Acad. Sci. USA 88:2432.*
12. McCafferty, J., A. D. Griffiths, G. Winter, and D. J. Chiswell. 1990. Phage antibodies: filamentous phage displaying antibody variable domains. *Nature 348:552.*
13. Kang, A. S., C. F. Barbas, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1991. Linkage of recognition and replication functions by assembling combinatorial antibody F(ab) libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA 88:4363.*
14. Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter. 1991. Making antibody fragments using phage display libraries. *Nature 352:624.*
16. Breitling, F., S. Dubel, T. Seehaus, I. Klewinghaus, and M. Little. 1991. A surface expression vector for antibody screening. *Gene 104:147.*
17. Chang, C. N., N. F. Landolfi, and C. Queen. 1991. Expression of antibody Fab domains on bacteriophage surfaces: potential use for antibody selection. *J. Immunol. 147:3010.*
18. Smith, G. P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science 228:1315.*
19. Parmley, S. F., and G. P. Smith. 1988. Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene 73:305.*
20. de la Cruz, V. F., A. A. Lal, and T. F. McCutchan. 1988. Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *J. Biol. Chem. 263:4318.*
21. Bass, S., R. Greene, and J. A. Wells. 1990. Hormone phage: an improved method for enrichment of variant proteins with altered binding properties. *Proteins 8:309.*
22. Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower. 1990. Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. USA 87:6378.*
23. Hellström, I., D. Horne, P. Linsley, J. P. Brown, V. Brankovan, and K. E. Hellström. 1986. Monoclonal mouse antibodies raised against human lung carcinoma. *Cancer Res. 46:3917.*
24. Hellström, I., P. L. Beaumier, and K. E. Hellström. 1986. Antitumor effects of L6, an IgG2a antibody that reacts with most human carcinomas. *Proc. Natl. Acad. Sci. USA 83:7059.*
25. Goodman, G. E., I. Hellström, L. Brodzinsky, C. Nicaise, B. Kulander, D. Hummel, and K. E. Hellström. 1990. Phase I trial of murine monoclonal antibody L6 in breast, colon, ovarian, and lung cancer. *J. Clin. Oncol. 8:1083.*
26. Fell, H. P., M. A. Gayle, D. Yelton, L. Lipsich, G. L. Schieven, J. S. Marken, A. Aruffo, K. E. Hellström, I. Hellström, and J. Bajorath. 1992. Chimeric L6 anti-tumor antibody: genomic construction, expression, and characterization of the antigen binding site. *J. Biol. Chem. In press.*
27. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt. Biochem. 162:156.*
28. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science 239:487.*
29. Sastry, L., M. Alting-Meese, W. D. Huse, J. M. Short, J. A. Sorge, B. N. Hay, K. D. Janda, S. J. Benkovic, and R. A. Lerner. 1989. Cloning of the immunological repertoire in *Escherichia coli* for generation of monoclonal catalytic antibodies: construction of a heavy chain variable region-specific cDNA library. *Proc. Natl. Acad. Sci. USA 86:5728.*
30. Huse, W. D. Combinatorial antibody expression libraries in filamentous phage. In *Antibody Engineering: A Practical Guide,* C. A. K. Borrebaeck, ed. W. H. Freeman and Company, New York, 1991, pp. 103–120.
31. Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA 82:488.*
32. Kunkel, T. A., J. Roberts, and R. Zakour. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol. 154:367.*
33. Hellström, K. E., D. E. Yelton, H. P. Fell, D. Beaton, M. Gayle, M. MacLean, M. Kahn, and I. Hellström. 1990. Epitope mapping and use of anti-idiotype antibodies to the L6 monoclonal anticarcinoma antibody. *Cancer Res. 50:2449.*
34. Russel, M., and P. Model. 1981. A mutation downstream from the signal peptidase cleavage site affects cleavage but not membrane insertion of phage coat protein. *Proc. Natl. Acad. Sci. USA 78:1717.*
35. Marks, J. D., H. R. Hoogenboom, T. P. Bonnert, J. McCafferty, A. D. Griffiths, and G. Winter. 1991. By-passing immunization: human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol. 222:581.*
36. Glaser, S. M., D. E. Yelton, and W. D. Huse. 1992. Antibody engineering by codon-based mutagenesis in a filamentous phage vector system. *J. Immunol. 149:3903.*

Pergamon
0161-5890(95)00098-4

Molecular Immunology, Vol. 32, No. 17/18, pp. 1387–1397, 1995

Printed in Great Britain
0161-5890/95 $9.50 + 0.00

A PHAGE DISPLAY SYSTEM FOR DETECTION OF T CELL RECEPTOR–ANTIGEN INTERACTIONS

TAKEBUMI ONDA,*†‡ DRAKE LAFACE,*‡§ GOTTFRIED BAIER,*‖ THOMAS BRUNNER,* NAKAYUKI HONMA,¶ TOSHIFUMI MIKAYAMA,¶ AMNON ALTMAN* and DOUGLAS R. GREEN*

*Division of Cellular Immunology, La Jolla Institute for Allergy and Immunology, 11149 N. Torrey Pines Rd, La Jolla, CA 92037, U.S.A.; ¶Kirin Brewery Co. Ltd, Pharmaceutical Research Laboratory, Maebashi Shi Gunma; Japan

(*First received* 29 *June* 1995; *accepted in revised form* 30 *June* 1995)

Abstract—The process of T cell recognition involves a complex set of interactions between the various components of the TCR/MHC peptide trimolecular complex. We have developed a system for exploring the specific binding interactions contributed by the constituent subunits of TCR complexes for components of their ligands. We utilized an M13 phage display system, designed for multivalent receptor display, to explore specific binding interactions between various TCRα chains and specific antigen in the absence of MHC. The multivalent TCR–phage display system was sensitive enough to reveal some TCRα chains capable of binding directly to antigen with the same fine specificity shown by the MHC-restricted T cells from which the α chains were derived. Cross-specificity analysis using two antigen-binding TCRα chains derived from T cells with different polypeptide antigen specificities confirmed the fidelity of this binding. In mixtures of antigen-binding and non-binding TCRα-displaying phage, specific selection was achieved at a starting frequency of 1/1000, suggesting that this system can be employed for selection and analysis of TCR-displaying phage libraries. While the binding specificities exhibited by these TCRs are unusual, they provide a novel perspective from which to study the specific binding interactions that constitute TCR antigen binding.

Key words: M13 filamentous phage, peptide binding, phage-display, TCR.

INTRODUCTION

Specific recognition by T lymphocytes is mediated by T cell antigen receptors (TCR) (Jorgensen *et al.*, 1992b; Chien and Davis, 1993). The most extensively studied TCR–ligand interactions involve TCRαβ heterodimers binding to peptide–MHC complexes. A number of studies have focused on defining the intricate biochemical interactions between the various components of the TCR and MHC–peptide complexes. As a consequence of these investigations, it has been suggested that residues within the CDR3 regions of TCRα and β chains interact with specific residues from the antigenic peptide fragment (Davis and Bjorkman, 1988; Engel and Hedrick, 1988; Danska *et al.*, 1990; Jorgensen *et al.*, 1992a). More recent studies have indicated that TCR interactions with peptide appear to be very specific, while interactions with MHC are comparatively degenerate (Ehrich *et al.*, 1993). The TCR–MHC interactions could be established in a variety of configurations with the same TCR and MHC and appear to be influenced by the interaction of the TCR with the peptide. Thus, while TCR interactions with residues from both peptide (Engel and Hedrick, 1988; Danska *et al.*, 1990; Jorgensen *et al.*, 1992a) and MHC (Ajitkumar *et al.*, 1988; Peccoud *et al.*, 1990) appear to be essential for efficient T cell recognition, the dominant interactions are seemingly mediated by CDR3 residues associated with specific peptide (Ehrich *et al.*, 1993).

Although most studies have focused on conventional TCR binding of MHC–peptide ligands, a number of alternative TCR–ligand interactions have emerged, providing the basis for further studies in TCR specificity. Examples include TCR β chain-binding to superantigens (White *et al.*, 1989; Gascoigne and Ames, 1991; White *et al.*, 1993); TCRs which specifically bind peptide-free mycolic acid antigens in the context of CD1 restricting elements (Porcelli *et al.*, 1992); and TCRs with specificities for carbohydrate moieties of post-translationally modified peptide antigens (Haurum *et al.*, 1994; Michaëlsson *et al.*, 1994). These recent studies of non-conventional TCR–ligand interactions lend credence to earlier studies of T cells which expressed αβ TCRs capable of binding directly to hapten-conjugates (Rao *et al.*,

†Current address: National Children's Medical Research Center, Tokyo, Japan.
‡These authors contributed equally and share first authorship.
§Author to whom correspondence should be addressed: Drake M. LaFace, La Jolla Institute for Allergy and Immunology, 11149 N. Torrey Pines Rd, La Jolla, CA 92037, U.S.A.
‖Current address: Institute for Medical Biology and Human Genetics, University of Innsbruck, Innsbruck, Austria.

1984a, 1984b; Siliciano *et al.*, 1986) or peptide antigens (Esch and Thomas, 1990) in the absence of MHC proteins. Interestingly, direct TCR-binding interactions with hapten or peptide ligands, in the absence of MHC, demonstrated relatively high affinity constants (K_Ds = 5×10^{-5}–6×10^{-6}) when compared to the normal range of TCR–ligand affinities (Rao *et al.*, 1984b; Siliciano *et al.*, 1986; Esch and Thomas, 1990; Davis and Chien, 1993).

Studies of TCR that directly bind nitrophenol haptens suggested that the specificity was mediated by the TCRα subunit in a manner independent of TCRβ (Kuchroo *et al.*, 1991). Similar studies by our group and others have suggested the possibility that a subset of TCRα molecules may bind directly to antigen for which the T cell encoding the TCR protein is specific (Bissonnette *et al.*, 1991; Green *et al.*, 1991; Mori *et al.*, 1993). While these unusual TCR–ligand interactions may represent a relatively small subset of TCR specificities, they are nonetheless intriguing because they provide an alternative approach with which to explore the many complex interactions that mediate specific binding. Since the constituent α and β chains of TCR may bind different residues of the peptide ligand component in a relatively autonomous manner (Jorgensen *et al.*, 1992a, 1992b), we reasoned that it should be feasible to study directly the specific interactions of TCRα chains with antigenic peptide. An elemental analysis of the specific interactions that constitute the complex binding of TCR with MHC–peptide ligands would be extremely useful. While these studies would not refute existing theories of conventional TCR-antigen binding, they would provide a unique perspective from which to study the individual specific binding interactions between the various subunits of the TCR–MHC/peptide complex which contribute to the overall TCR binding specificity.

We employed the filamentous phage display system for the study of immunological receptors (Kang *et al.*, 1991; Barbas and Lerner, 1991b) to directly assess specific binding interactions of several TCRα polypeptides with antigenic peptide and globular proteins in the absence of MHC. The multivalent phage display system, utilized to enhance avidity, was optimally suited to study interactions of TCR molecules directly with their peptide ligand, as the binding affinities are relatively low compared to immunoglobulin receptors.

EXPERIMENTAL PROCEDURES

Reagents, strains, vectors

The *Escherichia coli* XL1-Blue (tet^r) and VCSM13 helper phage (kan^r) were purchased from Stratagene. All enzymes were purchased from Boehringer Mannheim and Promega. The phagemid vector pComb8 (Kang *et al.*, 1991) utilized to produce clones encoding TCRα chains fused to the cpVIII of M13 filamentous phage was generously provided by Denise Burton (The Scripps Research Institute). All molecular biology procedures were performed according to conventional techniques as described in *Molecular Cloning: a Laboratory Manual* (Sambrook *et al.*, 1989) unless indicated otherwise.

Peptides and antibodies

Poly 18 peptides were kindly provided by Dr Bhagirath Singh (University of Western Ontario, Canada). Bee venom phospholipase A_2 and bovine phospholipase A_2 were purchased from Sigma. Cell lines producing hamster anti-mouse TCRα (H28-710.16) and anti-TCRβ (H57-597 hamster IgG) were kindly provided by Dr Ralph Kubo (Cytel, La Jolla, CA, U.S.A.). Supernatants were concentrated by precipitation with 45% saturated ammonium chloride followed by dialysis, and the antibodies were purified by protein A chromatography.

The TCRα cDNA and clones

Full-length cDNA sequences encoding the A1.1 TCRα (Green *et al.*, 1991), 5C.C7 TCRα (Fink *et al.*, 1986) (provided by S. Hedrick, University of California, San Diego, CA, U.S.A.) and the 3B3 TCRα (Mori *et al.*, 1993) (T. Mikayama, unpublished sequence), respectively, were used as templates in the polymerase chain reaction (PCR) subcloning procedures. Fifty nanograms of cDNA template was mixed with 60 pmol of each primer (Table 1), 200 mM dNTP, Promega *Taq* polymerase buffer containing 1.5 mM $MgCl_2$ and five units of *Taq* polymerase in a final volume of 100 ml. Amplification was carried out for 20–30 cycles on a TwinBlock thermal cycler (Ericomp Inc., San Diego, CA, U.S.A.) under the following conditions: 1 min of denaturation at 94°C, 2 min of primer annealing at 50°C, 1 min of elongation at 72°C, followed by a final elongation at 72°C for 15 min. Amplified products were hydrolysed (4 hr at 37°C) at the *Xho*I and *Xba*I sites (25 U enzyme/μg PCR fragment) encoded by the primers, size-fractionated by agarose gel electrophoresis and purified using a Geneclean procedure (Bio 101 Inc., La Jolla, CA, U.S.A.). The PCR primers were designed to produce TCRα cDNA sequences that could be ligated into the pComb8 vector in frame with the pelB leader sequence at the 5′ end and fused to the N-terminus of the cpVIII protein on the 3′ end. The purified *Xho*I – *Xba*I insert was directionally cloned into the phagemid vector pComb8 at the *Xho*I and *Spe*I sites and transformants were screened with internal oligonucleotide probes in colony lifts. Plasmid DNA from positive clones were sequenced by the dideoxy method using Sequenase 2.0 (USB Corp., Cleveland, OH, U.S.A.), analysed using the MacVector analysis program (IBI, New Haven, CT, U.S.A.) confirming identity, orientation and in-frame cloning of TCRα cDNAs.

Recombinant phage preparation

Production of recombinant filamentous phage displaying TCRα chains was carried out essentially as previously described for the production of phage displaying immunoglobulin receptors (Kang *et al.*, 1991; Barbas and Lerner, 1991b). Briefly, *E. coli* XL1-Blue (Stratagene) cells were transformed with recombinant TCRα/pComb8

Table 1. The TCRα-specific primers

Variable region primers	A1
TCR Vα (sense) 5C.C7	5′-GAAGAGCTCGAGATGAAATCC TGAGT-3′
TCR Vα (sense) 3B3	5′-TGGACACTCGAGATGCAGAGGAACC
TCR Vα (sense)	5′-GCGCCGCTCGAGATGAAATCCTTGAGTGTTTTACTA GTGGTCCTGTGGCTCCAGTTAAACTGCGTGAGGAGC- CAGCAGCAAGTGCAGCAGAGTCCTGCA-3′
Constant region primers	
TCR-VJC	
TCR Cα (antisense)	5′-GCTGTCTCTAGAGCCACCGCCACCGTCGACG-TACACAGCAGGTTCTGGGTT-3
TCR-VJ	
Truncated Cα (antisense)	5′-CAGGAGTCTAGAGCCACCGCCACCGTCGACG-TTTGAAAGTTTAGGTTCATATC-3′

Restriction sites used

phagemid and selected on LB plates containing 100 μg/ml ampicillin. Fresh ampicillin-resistant colonies were grown in liquid culture on a shaker at 37°C in super-broth (SB) for 1–2 hr in the presence of ampicillin (50 μg/ml) to select for cells bearing phagemid and tetracycline (10 μg/ml) to induce the F′ episome. The cultures were then incubated with 1 mM isopropyl β-D-thiogalactoside (IPTG) for 1 hr to induce production of the TCR–cpVIII fusion protein and then superinfected with VCSM13 helper phage (10^{12} pfu) (Stratagene) to give a final phage/cell ratio of 10–20:1. The cells were shaken (37°C) for an additional 2 hr and then selected for helper phage induced antibiotic resistance (kanamycin, 70 μg/ml) and grown overnight, 37°C, 250 rpm. Phage supernatant was cleared by centrifugation of the cultures (4000 rpm in a GSA rotor, Sorvall, at 4°C). The phage were precipitated by adding 3% (w/v) NaCl and 4% (w/v) polyethylene glycol 8000 for 1 hr at 4°C and centrifuged at 9000 rpm in a Sorvall GSA rotor at 4°C. The phage pellets were resuspended in PBS to 1/25 of the original volume and aggregated phage and debris were removed by centrifugation for 5 min in a benchtop microcentrifuge. Phage supernatants were transferred to fresh tubes and stored in aliquots at −20°C.

Affinity selection panning

The panning procedure utilized to screen binding of TCR displayed on the surface of filamentous phage was a modification of the original protocol described by Parmley and Smith (1988). Wells of a microtitration ELISA plate (Immulon 2, Dynatech) were coated with peptide or phospholipase A_2 (Sigma) in sodium bicarbonate (0.1 M, pH 9.5) at 4°C overnight in a volume of 50 μl at the concentrations indicated in the figure legends. Antibodies were coated onto plates using Tris (50 mM, pH 9) at 4°C overnight in a volume of 50 μl at the concentrations indicated in the figure legends. The wells were washed twice with PBS (Dulbecco's, pH 7.4) and blocked by completely filling the well with 3% bovine serum albumin (BSA) (Sigma, fraction V) and incubated for 1 hr at 37°C. The blocking solution was flicked out, rinsed with PBS and 50 μl recombinant phage was added to each well (typically 10^{10}–10^{11} CFU unless otherwise indicated) and incubated for 2 hr at 37°C. The phage were then removed and the wells were washed 10× (pipetting up and down to wash) with TBS/Tween (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 0.05% Tween 20) over a period of 1 hr allowing the wells to remain filled between washings. Washing was done carefully to ensure similarity between samples, experiments and individuals. Adherent phages were eluted with 50 μl elution buffer (0.1 M HCl adjusted to Ph 2.2 with solid glycine, 1 mg/ml BSA) for 10 min at room temperature and neutralized with 6 ml of 1 M Tris base. Eluted phage were used to infect fresh *E. coli* XL1-Blue cells, OD_{600} = 1 (grown in the presence of 10 μg/ml tetracycline to induce F′ episome expression) for 15 min at 37°C followed by selection on LB/ampicillin plates to assess the number of bound recombinant phage per ml of eluted phage.

Immunoblot of TCRα–pVIII fusion proteins

Phagemid DNA of TCR-pComb8 or control constructs were transformed in XL1-Blue cells and transformants were inoculated in 25 ml 2×TY medium containing 50 μg/ml ampicillin. Protein expression was induced with 0.1 mM IPTG at an OD_{600} = 1 for 7 hr at 26°C. Cells were harvested by centrifugation, sonicated on ice for 1 min in PBS, 1 mM PMSF, 1% NP-40 and the resulting lysate was centrifuged (13,000g, 4°C) to remove insoluble debris. The soluble cellular proteins were analyzed by SDS–PAGE and immunoblotted with an anti-TCRα antibody (H28-710.16).

Presentation assay

The TCR specificity of A1.1 hybridoma cells for our poly 18-related peptide analogues was assessed by measuring the IL-2 responses of A1.1 cells to peptide-pulsed BALB/c ($I\text{-}A^d$) spleen cells (γ-irradiated 2000 rad). Supernatants were collected following overnight peptide stimulation and the relative production of IL-2 was assessed in a CTLL assay. Starved CTLL cells were washed and

incubated in various dilutions of sample or control supernatants (100 μl) in 96-well plates (Costar) at 37°C for 16 hr. The cells were then pulsed with [^{3}H]-thymidine (1 μCi/well) and harvested an additional 4–6 hr later. The IL-2-dependent proliferation was assessed by the relative level of [^{3}H]-thymidine incorporation (cpm).

Hybridization of replica plates

Following two rounds of affinity selection panning (described above), ampicillin-resistant CFUs were blotted onto replica filters as described (Sambrook *et al.*, 1989). The replica filters were denatured (0.5 N NaOH, 1.5 M NaCl) and neutralized [1.5 M NaCl, 0.5 M Tris–HCL (pH 7.4)] and then fixed (80°C, 2 hr). The filters were then hybridized with a 5′ ^{32}P-labelled A1.1 TCR Vα specific probe (5′-GAAGAGCTCGAGATGAAATC-CTTGAGT-3′).

RESULTS

The TCR phage display

We previously presented evidence suggesting that the TCRα molecule from the A1.1 T cell hybridoma (Vα1.2, JαTA65) may bind directly to antigenic peptides for which this T cell is specific (Bissonnette *et al.*, 1991; Green *et al.*, 1991). The M13 filamentous phage display system (Kang *et al.*, 1991; Barbas *et al.*, 1991a; Barbas and Lerner, 1991b) provided the means to directly assess the ability of A1.1 TCRα polypeptides to specifically bind to peptide antigens. We utilized the phagemid vector, pComb8 (Kang *et al.*, 1991) to generate recombinant phage-displaying TCR proteins fused to the N-terminus of the coat protein VIII (cpVIII). The cpVIII molecules form the capsid coat during phage assembly (utilizing about 2500 cpVIII proteins per phage particle) allowing for multiple receptors to be displayed on the surface of the recombinant phage (Felici *et al.*, 1991; Greenwood *et al.*, 1991; Kang *et al.*, 1991). This multivalent expression can greatly enhance the avidity, which facilitates the study of binding interactions with moderate affinity.

The PCR-generated fragments encoding the VJ- or VJC-domains of the A1.1 TCRα molecule were subcloned into the pComb8 phagemid vector to generate recombinant phage displaying multiple copies of the A1.1 TCRα molecule on the surface. This is illustrated in Fig. 1 (see Experimental procedures). Extracts of cells transformed with various recombinant or control phagemids were subjected to SDS–PAGE and immunoblotted with an anti-TCR Cα specific antibody (H28.710.16) to confirm that TCRα/cpVIII fusion proteins were produced. As shown in Fig. 2A, a protein that migrated with the predicted molecular weight for the TCR–VJC/cpVIII fusion protein, Mr 36.5 kDa, was detected from extracts of cells transformed with the A1.1 TCR–VJCα construct. A second protein of M_r 34.5 kDa was also detected, which probably represents a partially degraded form of the fusion protein, since it was not observed in the other lanes. No bands were observed from extracts of cells transformed with either the parental phagemid (pComb8, vector only) or the A1.1 TCR VJα construct, which both lack the Cα domain recognized by the H28.710.16 antibody used for detection.

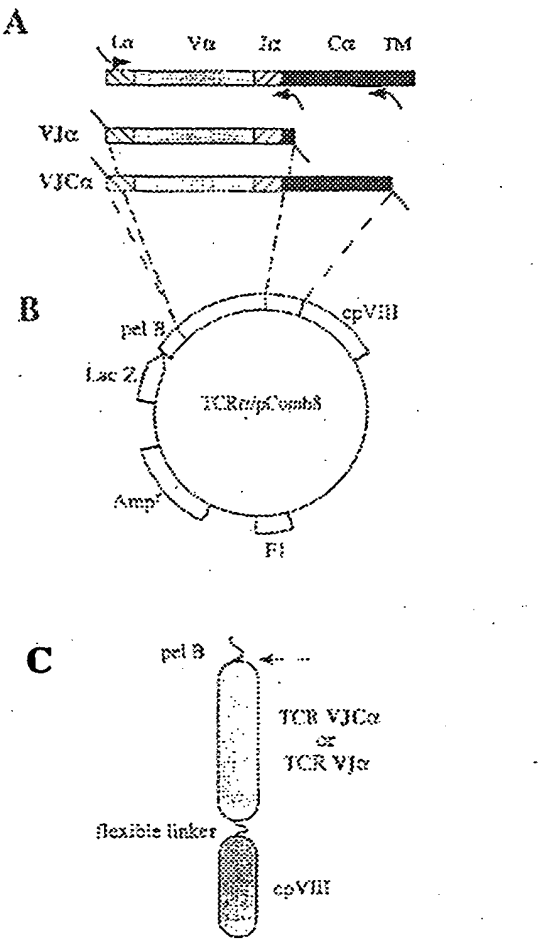


Fig. 1. Schematic protocol for the generation of chimeric constructs. (A) Full-length cDNA sequences encoding various TCRα chains were used as templates in the polymerase chain reaction (PCR) subcloning procedures. The 5′ PCR primers (arrows) contained an *Xho*I site designed to generate a PCR fragment that could be ligated into the pComb8 vector in frame with the pelB leader sequence. The 3′ primers (arrows) contained a sequence encoding a flexible linker (GGGS) and an *Xba*I site to generate a fusion gene with the major cpVIII upon ligation. The PCR fragments were digested with *Xba*I and *Xho*I, purified and directionally cloned into the phagemid vector pComb8. (B) The recombinant TCRα–pComb8 phagemids contained an amp^r gene utilized to facilitate drug selection of recombinant phage. Production of the recombinant TCRα/cpVIII fusion protein was driven by the inducible LacZ promoter following transformation of competent XL1-Blue cells (Stratagene). Transport into the periplasmic space during production was mediated by the pelB leader sequence. (C) The predicted structure of the recombinant TCRα/cpVIII fusion proteins consisted of a TCRα polypeptide and a flexible linker (GGGS) fused to the N-terminus of the cpVIII molecule. The pelB leader sequence at the N-terminal end is cleaved off following transport into the periplasmic space. The recombinant fusion protein was incorporated into the viral capsid coat during phage assembly to generate recombinant M13 phage displaying TCR molecules on the surface.

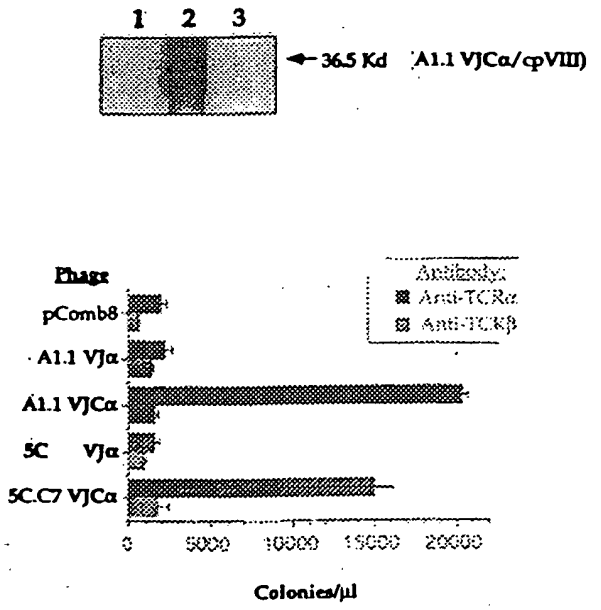


Fig. 2. Expression of TCRα/pComb8 recombinant fusion proteins. (A) Western blot analysis to detect expression of A1.1-TCR–VJCα/cpVIII fusion proteins in *E. coli*. Lysates from recombinant TCRα–pComb8 (lane 1: A1.1-TCR VJ; lane 2: A1.1-TCR–VJC) or parental pComb8 (lane 3: vector only) phagemid transformed XL1-Blue cells were subjected to SDS–PAGE and immunoblotted with an anti-TCR Cα antibody (H28.710.16). The position of the band migrating at the predicted molecular weight of the TCR–VJCα/cpVIII fusion protein (M, 36.5 kD) is indicated by the arrow. (B) Display of TCR–VJCα polypeptides on the surface of recombinant phage. Binding of recombinant phage to anti-TCR antibody was utilized to detect the presence of TCR on the phage surface. Recombinant TCRα–pComb8 phage or parental pComb8 phage (indicated in the figure) were panned on plates coated with antibodies specific for TCR Cα (H28.710.16, solid bars) or TCR Vβ6 (H57-597, hatched bars). The number of bound phage was assessed by counting the number of eluted ampicillin resistant CFU following extensive washing as described in the experimental procedures.

The presence of TCRα molecules displayed on the surface of recombinant phage was assessed using anti-TCR antibodies. Recombinant phage were "panned" on plastic dishes (Parmley and Smith, 1988) coated with either anti-TCR Cα (H28-710.16) or anti-TCR Cβ (H57-597) antibodies (Fig. 2B). Following extensive washing, the bound phage were eluted with a low pH buffer. Subsequent infection of host cells with eluted phage was readily achievable as the cpIII proteins, utilized for adsorption to the F′ episome and infection of *E. coli*, remain intact following acid elution. Since the pComb8 phagemid vector contains an *amp*^{r} gene, the number of bound phage could be assessed by counting the number of ampicillin-resistant CFU following infection of host cells with eluted phage. In addition to recombinant phage displaying A1.1 TCRα chains, we also constructed phage displaying the TCRα chains from the 5C.C7 T cell hybridoma (specific for pigeon cytochrome C peptide restricted to I-E^k) (Fink *et al.*, 1986). As shown in Fig. 2B, both the A1.1 TCR–VJCα as well as the 5C.C7-TCR–VJCα recombinant phage were readily bound by the anti-TCRα but not the anti-TCRβ antibodies. As predicted, the parental pComb8 phage (vector only) and recombinant TCR–VJα phage showed no binding to either antibody as these phage lack the TCR Cα determinant recognized by H28.710.16 antibody. Thus, polypeptides bearing TCR Cα specific determinants were effectively displayed on the surface of recombinant phage particles, indicating that TCRα/cpVIII fusion proteins could be effectively incorporated into the capsid coat during phage assembly.

Specificity of TCR displayed on recombinant phage

The A1.1 T hybridoma cells recognize a synthetic polypeptide with the sequence poly [EYK(EYA)$_5$] (poly 18) presented by I-A^d (Fotedar *et al.*, 1985). The antigenic fine specificity of A1.1 cells for a series of poly 18-related peptide analogues was assessed by measuring the IL-2 responses to peptide pulsed BALB/c (I-A^d) spleen cells. As shown in Fig. 3A, A1.1 cells produced cytokine in response to two of the peptides, EYK(EYA)$_4$ and EYK-(EYA)$_4$EYK. Peptide analogues, substituted with alanines at residues 3 or 10, failed to stimulate A1.1 cells. These results are in agreement with those reported by others for this cell line (Fotedar *et al.*, 1985). The antigenic fine specificity of the A1.1 hybridoma cells for the four peptides described above provided a model with which to compare specificity of TCRα molecules displayed on recombinant phage.

The poly 18-related peptide analogues utilized to characterize the specificity of A1.1 hybridomas were coated onto 96-well plates for panning experiments to assess the binding capacity and specificity of the TCRα molecules displayed on the phage surfaces. Bound phage were extensively washed, eluted and then quantitated by counting the number of ampicillin-resistant CFUs following infection of host cells. Representative of a series of 10 binding experiments, Fig. 3B shows that recombinant phage expressing A1.1 TCR–VJCα preferentially bound to the antigenic peptides, EYKEYAEYAEYAEYAEYK and EYKEYAEYAEYAEYA, but not to non-antigenic peptides with an alanine substitution at residues 3 (EY*A*-EYAEYAEYAEYA) or 10 (EYKEYAEYA*A*YAEYA). The parental pComb8 phage (vector only) and recombinant phage displaying 5C.C7-TCR–VJCα showed no specific binding to any of the four poly 18 peptide analogues (Figs 3B and 4A). Recombinant phages displaying A1.1 TCR VJα, which lack the Cα-domain, also showed preferential binding to the antigenic peptides, indicating that the specific binding was mediated by the TCR Vα- and Jα-region domains (Fig. 4B). The observed preferential binding to antigenic peptide was reproducible and consistent over a wide range of phage titres when we collected the quantitative data from seven different experiments utilizing several different recombinant phage preparations from two different recombinant constructs (Fig. 4C). Thus, direct binding by phage displaying A1.1 TCRα to peptides reflects the antigenic specificity of the

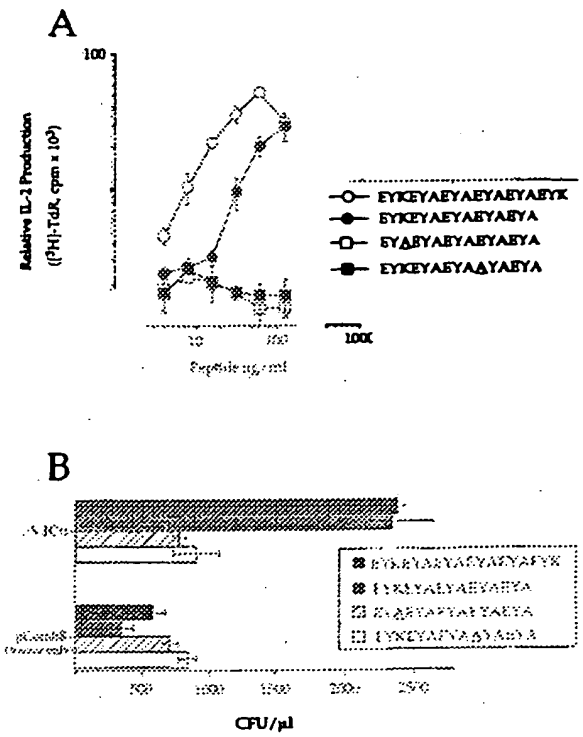


Fig. 3. Antigenic fine specificity of A1.1 hybridoma cells and A1.1-TCRα Recombinant phage. (A) Fine specificity of A1.1 hybridoma cells. The ability of poly 18 related peptide analogues to induce an IL-2 response was utilized to assess the fine specificity of A1.1 hybridoma cells. The BALB/c spleen cells (I-A^d) were pulsed with the indicated poly 18 related peptide analogues and co-cultured with A1.1 hybridoma cells as described in the experimental procedures. The relative amount of IL-2 produced in response to the peptide antigens is indicated in the *Y*-axis as the amount of tritiated thymidine uptake in a CTLL assay. (B) Comparative fine specificity of peptide binding by recombinant A1.1-TCR phage. The binding capacity of recombinant phage displaying A1.1-TCR–VJCα was assessed using the same four poly 18 related peptide analogues used to assess the fine specificity of the A1.1 hybridoma cells. Plates were coated with the indicated peptide analogues followed by panning with recombinant phage displaying A1.1 TCR–VJCα. The number of bound phage following extensive washing was assessed by determining the number of eluted ampicillin resistant CFU.

A1.1 hybridoma from which the TCRα cDNA sequence was derived (see Fig. 3A), despite the absence of class II MHC presentation.

Affinity selection of specific TCRα

To assess the capacity for affinity selection of phage displaying specific TCR polypeptides, recombinant phage displaying A1.1 TCR–VJCα were mixed 1:100 or 1:1000 with phage displaying 5C.C7-TCR–VJCα and then subjected to two rounds of affinity enrichment on plates coated with antigenic poly 18 peptide analogues. The eluted phage were used to infect host cells and colonies from ampicillin plates were lifted onto nitrocellulose filters. The colony-lifts were lysed and hybridized with a DNA probe specific for the A1.1 TCRα sequence. We found that 85% (468/554) of the colonies from the 1:100 mix and 76% (113/149) of the colonies from the 1:1000 mix contained A1.1 TCRα specific sequences (Fig. 5). Thus, two rounds of selection enriched the phage displaying A1.1 TCRα by over 500- and 3000-fold, respectively. These results suggest that the phage-display system can be utilized to select for T cell receptors capable of binding directly to peptide ligands and has potential to be utilized as a gene identifying system for TCRα chains with relatively high affinity for their peptide ligand.

The TCRα binding directly to protein antigen

While phage displaying 5C.C7-TCR–VJCα failed to bind to the poly 18 peptide analogues bound by A1.1 TCRα recombinant phage, these phage also failed to bind antigenic cytochrome C peptide, in repeated attempts (data not shown). Thus, differences in detectable binding of TCR Vα to specific antigenic peptides may represent variance in the particular TCR–ligand interactions such that only a subset of TCR, including that on A1.1 cells, have direct interactions strong enough to be detectable in this system. Accordingly, we constructed recombinant phage displaying TCRα from the 3B3 T cell hybridoma, which has previously been suggested to express a TCRα chain product capable of direct interaction with antigen (Mori *et al.*, 1993). The 3B3 hybridoma was generated from BALB/c mice immunized with bee venom phospholipase A_2 (PLA_2) and is specific for a bee venom PLA_2 derived peptide (residues 19–34) in the context of I-A^d (Mori *et al.*, 1993). Two aspects of 3B3 TCR binding characteristics made it particularly interesting for phage display studies of direct antigen-binding TCRα chains. Firstly, three-dimensional structural analysis of bee venom PLA_2 indicated that the antigenic peptide residues associated with 3B3 TCR binding are exposed to the outer surface of the PLA_2 molecule in the form of a flexible loop (Scott *et al.*, 1990). Secondly, previous peptide competition studies of 3B3 TCR specificity suggested that the TCRα chain may be able to bind directly to unprocessed bee venom PLA_2 (Mori *et al.*, 1993). Therefore, we constructed recombinant phage displaying 3B3-TCR–VJCα to test the ability of this TCRα chain to directly bind to unprocessed bee venom PLA_2 protein in a specific manner.

The PCR-generated fragments encoding the 3B3-TCR–VJCα polypeptide were subcloned into the pComb8 vector for production of recombinant phage. Bee venom PLA_2 was coated onto 96-well plates for panning experiments to assess the binding capacity and specificity of the 3B3 TCRα chains displayed on the phage surface. We observed binding of recombinant 3B3-TCR–VJCα phage, but not parental pComb8 phage (vector only), to bee venom PLA_2 coated wells (Fig. 6A). Furthermore, binding of recombinant 3B3-TCR–VJCα phage to BSA (3% BSA, used for blocking) was not observed. Moreover, the recombinant phage displaying 3B3 TCR–VJCα did not bind to bovine PLA_2 (Fig. 6B), which lacks the exposed hydrophilic loop containing the

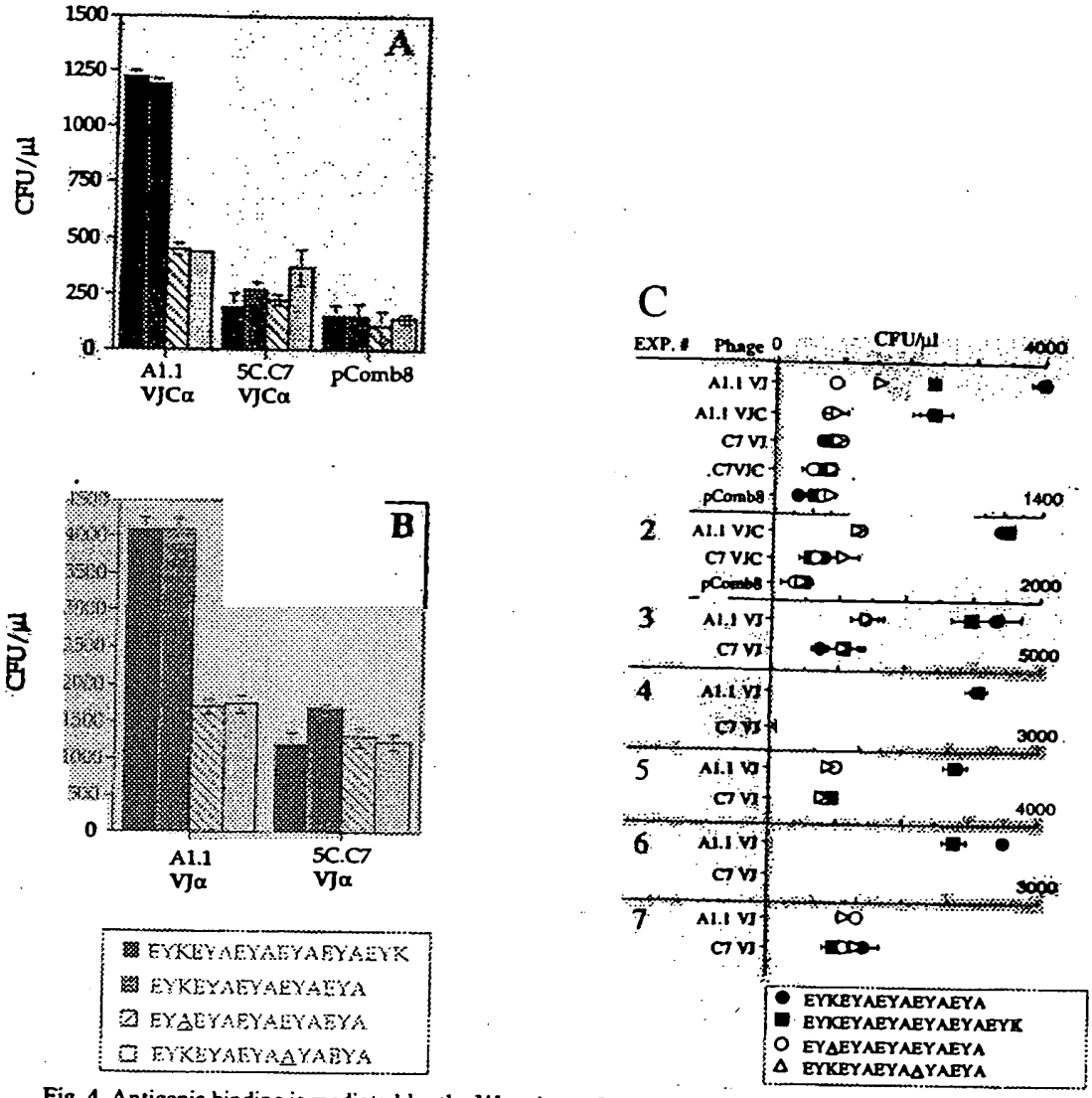


Fig. 4. Antigenic binding is mediated by the VJ regions of specific TCR-recombinant phage. (A) The binding capacity of parental pComb8 phage or recombinant phage displaying A1.1-TCR–VJCα or 5C.C7-TCR–VJCα for the indicated poly 18 related peptide analogues was assessed using the panning protocol described in the experimental procedures. Plates were coated with the indicated peptide analogues and panned with the indicated phage. The number of bound phage following extensive washing was assessed by determining the number of ampicillin resistant CFU/μl. (B) The binding capacity of recombinant phage displaying TCR VJα (truncated C-region) from the A1.1 or 5C.C7 hybridoma cells to the four indicated poly 18 related peptide analogues was determined as described above to assess whether the VJ-regions were sufficient to mediate the specific binding to antigenic peptides. (C) The relative binding specificity of several recombinant phage preparations, displaying TCR–VJCα or TCR VJα, for the indicated poly 18 related peptide analogues are compared from seven different panning experiments as described above. The experiments described in (A) and (B) are included for comparison.

antigenic peptide sequence associated with 3B3 TCR binding (Dijkstra *et al.*, 1981; Mori *et al.*, 1993). The observed preferential binding to antigenic bee venom PLA_2 was reproducible and consistent over a wide range of phage titres when we collected the quantitative data from five different experiments utilizing several different recombinant phage preparations (Fig. 6C). Thus, the phage display system enabled us to assess the novel capacity of a TCRα chain to bind directly to antigenic protein in the absence of MHC.

In order to further assess the specificity of the TCRα recombinant phages for their respective antigenic ligands, we did parallel panning experiments with the A1.1 TCR–VJCα and 3B3 TCR–VJCα recombinant phage (Fig. 7). Phage displaying A1.1 TCR–VJCα showed binding to the antigenic poly 18 peptide analogue (EYK-EYAEYAEYAEYAEYK) but not to bee venom PLA_2. In contrast, phage displaying 3B3 TCR–VJCα demonstrated specific binding to bee venom PLA_2 but not to the poly 18 peptide analogue. Thus, the TCR phage-

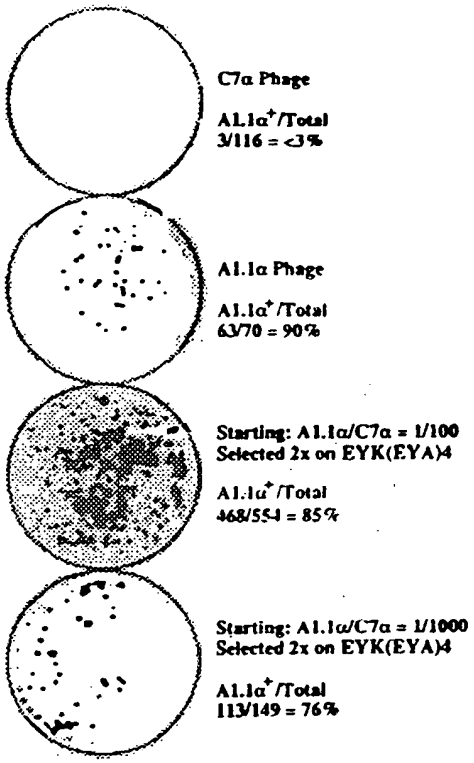


Fig. 5. Affinity selection of specific TCR on recombinant phage. A mixing experiment was done to assess the capacity for affinity selection of peptide specific TCR displayed on the surface of recombinant A1.1-TCRα–pComb8 phage. The A1.1-TCR–VJCα phages were mixed 1:100 or 1:1000 with 5C.C7-TCR–VJCα phage and subjected to two rounds of affinity selection by panning as described in the experimental procedures. Colony lifts were hybridized with a radiolabeled A1.1-TCRα specific oligonucleotide probe. Autoradiographs of the hybridized replica filters are shown in panels A–D. Autoradiographs of control colony lifts from plates inoculated with unmixed 5C.C7 TCRα recombinant phage (panel A) or unmixed A1.1-TCRα recombinant phage (panel B) are shown for comparison with plates from mixing experiments starting with A1.1α/5C.C7α ratios of 1:100 (panel C) or 1:1000 (panel D) prior to the two rounds of affinity selection.

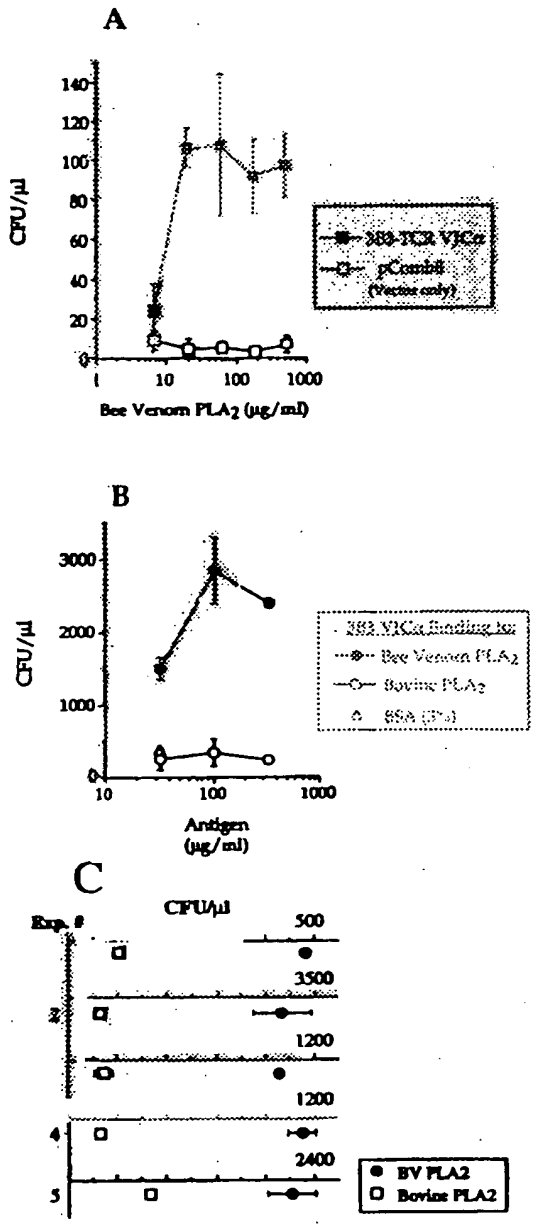


Fig. 6. Antigenic specificity of recombinant phage displaying TCRα polypeptide from 3B3 hybridoma cells. (A) The binding capacity of parental pComb8 phage or recombinant phage displaying TCR–VJCα from the 3B3 hybridoma cell to bee venom PLA_2 was assessed using the panning protocol described in the experimental procedures. Plates were coated with the indicated concentration of bee venom PLA_2 and then panned with 10^8 CFU of recombinant phage displaying 3B3-TCR–VJCα or parental pComb8 phage. The number of bound phage following extensive washing was assessed by determining the number of ampicillin resistant CFU/μl. (B) The specificity of recombinant phage displaying TCR–VJCα from the 3B3 hybridoma cell for bee venom PLA_2 and bovine PLA_2 was assessed using the panning protocol. The number of bound phage following panning with 10^{10} CFU of recombinant phage and extensive washing was assessed by determining the number of ampicillin resistant CFU/μl. (C) The relative binding specificity of several recombinant phage preparations, displaying 3B3-TCR–VJCα for bee venom PLA_2 or bovine PLA_2 are compared from five different panning experiments as described above.

display system described here facilitated the study of TCRα chains capable of binding directly to antigenic epitopes in a specific manner.

DISCUSSION

The phage-display system (Smith, 1985) has proved to be extremely useful for the study of specific receptor–ligand interactions (Winter *et al.*, 1994). The advantage of this system for immunological studies is that the display of receptors on recombinant phage provides a means to link directly antigen recognition structures and the genetic instructions encoding the receptor (Kang *et al.*, 1991). Consequently, the phage display system has been invaluable for the characterization and selection of antibody specificities for antigenic ligands. Here, we have shown that this system can be utilized further to study specific binding interactions of constituent TCR poly-

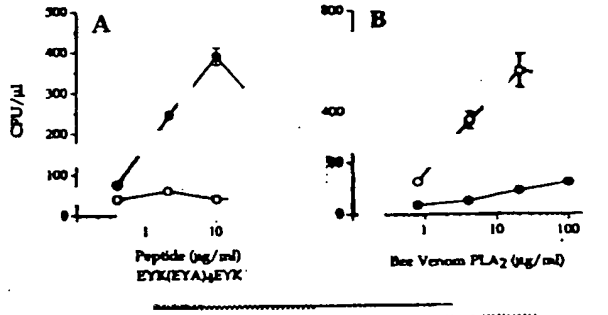


Fig. 2. Cross-cross experiments to assess relative specificity of recombinant phage. Recombinant phage displaying TCR-VJCα from the A1.1 or 3B3 hybridoma cells were assessed for their ability to bind the antigenic poly-18 related peptide analogue EYKEYAEYAEYAEYAEYK (A) or bee venom PLA2 (B) at various concentrations indicated in the figure, using the panning protocol described in the experimental procedures. The number of bound phage following extensive washing was assessed by determining the number of ampicillin resistant CFU/µl.

peptides to antigenic ligands. Specifically, the phage-display system facilitated a novel demonstration of direct binding by TCRα chains to peptide or protein antigens in a specific manner. These results indicate that for some TCRα chains the binding affinity to antigenic peptide may be sufficiently high that it can be detected in the absence of the TCRβ chain, and in an MHC-independent manner. It has been postulated previously that TCR binding to peptide by Vα- and Vβ-subunits may be relatively autonomous and that the resulting TCR–peptide interactions can dramatically influence TCR–MHC interactions; indicating the "primacy" of TCR–peptide interactions (Ehrich *et al.*, 1993). Our results extend these findings by demonstrating that the dominant interactions of certain TCRα chains for peptide antigens may be sufficiently high that they can be analysed independently. However, these interactions are quite unusual in that they do not require the expression of the second TCR subunit or normal MHC and coreceptor interactions. These results raise the concern that this model does not reflect typical TCR–ligand interactions. Indeed, gene transfer studies of A1.1 TCRα chain sequences indicated that the expression of the α chain alone was not sufficient to confer responsiveness to peptide antigen alone or in the context of I-A^d (H. Zheng, A. Fotedar and D.R. Green; unpublished data). Furthermore, activation of A1.1 hybridoma cells required specific peptide antigen to be presented in the context of the appropriate MHC (I-A^d) (Fotedar *et al.*, 1985). The requirement for MHC presentation and natural TCRαβ conformation for functional interaction of cell surface expressed receptors may be essential for bringing the respective binding sites into close enough proximity to establish the specific binding interaction between the residues of opposite charge which mediate binding. The specificity data presented here (Fig. 3 and Fig. 4 and Fig. 7) suggest that these requirements may not be necessary for some receptors to detect specific binding in non-cellular TCR phage-display systems. Therefore, the data presented here do not suggest that the TCRα chain alone would be sufficient for T cell recognition and activation, but rather that the TCR phage-display system provides a means to explore the specific binding interactions that mediate TCR binding to antigen at a constituent level. While this study does not refute commonly held theories of TCR–antigen interactions, this system does provide a means to explore the innate specific binding interactions contributed by the constituent subunits of TCR complexes for the peptide component of MHC–peptide ligands. In the absence of detailed structural data for the TCR/peptide/MHC trimolecular complex, models such as this can provide important insight into the nature of the elemental interactions that constitute TCR binding to MHC–peptide complexes. Furthermore, such information should facilitate interpretation of the structural data for the trimolecular complexes when this becomes available.

Previous studies indicate that TCR–peptide binding interactions are very specific, since single amino acid changes had profound effects on binding (Engel and Hedrick, 1988; Danska *et al.*, 1990). In accordance with these results, the direct binding of phage-displayed TCRα chains to antigen appeared to be specific. Recombinant phage-displaying A1.1 TCRα chains bound to antigenic peptide analogues but not to mutated peptide analogues that failed to activate A1.1 hybridoma cells (Fig. 3 and Fig. 4). Additionally, the cross-specificity experiment indicated that recombinant phage-displaying A1.1 TCR–VJCα or 3B3-TCR–VJCα bound their respective antigens but did not cross-react (Fig. 7).

The failure of repeated attempts to demonstrate specific binding of recombinant phage-displaying 5C.C7 TCRα to cytochrome C peptide indicated that only a subset of TCR Vα have the capacity for direct interactions with antigen strong enough to be detectable in this system. Previous studies of 5C.C7 TCR interactions with peptide antigen indicate that specific binding was mediated by CDR3 residues from both the α and β chains (Jorgensen *et al.*, 1992a, 1992b). Thus, binding to antigen by receptors such as 5C.C7 TCR may require the combined binding interactions of the α and β chains to mediate strong specific binding, while antigen binding by receptors such as A1.1 and 3B3 TCR may be dominated by specificities mediated predominantly by the TCRα chain.

The phage-display system described here extends the study of T cell recognition by providing a means to simplify the study of the specific interactions that constitute the overall receptor–ligand interactions. The power of the phage-display system is that recombinant phage can be amplified and subjected to multiple rounds of affinity selection as they contain the genetic instructions for production of the receptor molecule. A system for selection and cloning of antigen-binding TCR molecules will be useful in exploring TCR binding to various antigens and should provide a means to more closely examine TCR–ligand interactions.

Acknowledgements—We thank Dr K. Ishizaka and Dr A. Kang for helpful discussions. We thank Dr N. Gascoigne for critically reviewing the manuscript. We also thank Dr D. Burton (The Scripps Research Foundation) for the pComb8 phage display expression vector and Dr S. Hedrick (University of California, San Diego, U.S.A.) for the 5C.C7 TCRα cDNA sequence. This work was supported in part by National Institutes of Health grant GM52735. D.M.L. was a Myasthenia Gravis Foundation Fellow.

REFERENCES

Ajitkumar P., Geier S. S., Kesari K. V., Borriello F., Nakagawa M., Bluestone J., Saper M. A., Wiley D. C. and Nathenson S. G. (1988) Evidence that multiple residues on both the α helices of the class I MHC molecule are simultaneously recognized by the T cell receptor. *Cell* **54**, 47–56.

Barbas C. F., Kang A. S., Lerner R. A. and Benkovic S. J. (1991a) Assembly of combinatorial antibody libraries on phage surfaces: The gene III site. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7978–7982.

Barbas C. F. and Lerner R. A. (1991b) Combinatorial immunoglobulin libraries on the surface of phage (Phabs): Rapid selection of antigen-specific Fabs. *Meth. Compan. Meth. Enzymol.* **2**, 119–124.

Bissonnette R., Zheng H., Kubo R. T., Singh B. and Green D. R. (1991) A T helper cell hybridoma produces an antigen-specific regulatory activity. *J. Immun.* **146**, 2898–2907.

Chien Y.-H. and Davis M. M. (1993) How $\alpha\beta$ T-cell receptors 'see' peptide/MHC complexes. *Immun. Today* **14**, 597–602.

Danska J. S., Livingstone A. M., Paragas V., Ishihara T. and Fathman C. G. (1990) The presumptive CDR3 regions of both T cell receptor α and β chains determine T cell specificity for myoglobin peptides. *J. exp. Med.* **172**, 27–33.

Davis M. and Chien Y.-H. (1993) Topology and affinity of T-cell receptor mediated recognition of peptide–MHC complexes. *Curr. Opin. Immun.* **5**, 45–49.

Davis M. M. and Bjorkman P. J. (1988) T-cell antigen receptor genes and T-cell recognition. *Nature* **334**, 395–402.

Dijkstra B. W., Kalk K. H., Hol W. G. J. and Drenth J. (1981) Structure of bovine pancreatic phospholipase A_2 at 1.7 Å resolution. *J. molec. Biol.* **147**, 97–123.

Ehrich E. W., Devaux B., Rock E. P., Jorgensen J. L., Davis M. M. and Chien Y. H. (1993) T cell receptor interaction with peptide/major histocompatibility complex (MHC) and superantigen/MHC ligands is dominated by antigen. *J. exp. Med.* **178**, 713–722.

Engel I. and Hedrick S. M. (1988) Site-directed mutations in the VDJ junctional region of a T cell receptor β chain cause changes in antigenic peptide recognition. *Cell* **54**, 473–484.

Esch T. R. and Thomas D. W. (1990) Ia-independent binding of peptide antigen to the T cell receptor. *J. Immun.* **145**, 3972–3978.

Felici F., Castagnoli L., Musacchio A., Jappelli R. and Cesareni G. (1991) Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. *J. molec. Biol.* **222**, 301–310.

Fink P. J., Matis L. A., McKlligott D. L., Bookman M. and Hedrick S. M. (1986) Correlations between T-cell specificity and the structure of the antigen receptor. *Nature* **321**, 219–226.

Fotedar A., Boyer M., Smart W., Widtman J., Fraga E. and Singh B. (1985) Fine specificity of antigen recognition by T cell hybridoma clones specific for poly-18: a synthetic polypeptide antigen of defined sequence and conformation. *J. Immun.* **135**, 3028–3033.

Gascoigne N. R. and Ames K. T. (1991) Direct binding of secreted T-cell receptor beta chain to superantigen associated with class II major histocompatibility complex protein. *Proc. natn. Acad. Sci. U.S.A.* **88**, 613–616.

Green D. R., Bissonnette R., Zheng H., Onda T., Echeverri F., Mogil R. J., Steele J. K., Voralia M. and Fotedar A. (1991) Immunoregulatory activity of the T-cell receptor α chain demonstrated by retroviral gene transfer. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8475–8479.

Greenwood J., Willis A. E. and Perham R. N. (1991) Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from *Plasmodium falciparum* circumsporozoite protein as antigens. *J. molec. Biol.* **220**, 821–827.

Haurum J. S., Arsequell G., Lellouch A. C., Wong S. Y. C., Dwek R. A., McMichael A. J. and Elliot T. (1994) Recognition of carbohydrate by major histocompatibility complex class I-restricted glycopeptide-specific cytotoxic T lymphocytes. *J. exp. Med.* **180**, 739–744.

Jorgensen J. L., Esser U., Groth B. F. d. S., Reay P. A. and Davis M. M. (1992a) Mapping T-cell receptor–peptide contacts by variant peptide immunization of single-chain transgenics. *Nature* **355**, 224–230.

Jorgensen J. L., Reay P. A., Ehrich E. W. and Davis M. M. (1992b) Molecular components of T-cell recognition. *A. Rev. Immun.* **10**, 835–873.

Kang A. S., Barbas C. F., Janda K. D., Benkovic S. J. and Lerner R. A. (1991) Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4363–4366.

Kuchroo V. K., Byrne M. C., Atsumi Y., Greenfield E., Connolly J. B., Whitters M. J. Jr, R. M. O. H. Collins M. and Dorf M. E. (1991) T-cell receptor α chain plays a critical role in antigen-specific suppressor cell function. *Proc. natn. Acad. Sci. U.S.A.* **88**, 8700–8704.

Michaëlsson E., Malmström V., Reis S., Engström Å., Burkhardt H. and Holmdahl R. (1994) T cell recognition of carbohydrates on type II collagen. *J. exp. Med.* **180**, 745–749.

Mori A., Thomas P., Tagaya Y., Iijima H., Grey H. and Ishizaka K. (1993) Epitope specificity of bee venom phospholipase A_2-specific suppressor T cells which produce antigen-binding glycosylation inhibiting factor. *Int. Immun.* **5**, 833–842.

Parmley S. F. and Smith G. P. (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene* **73**, 305–318.

Peccoud J., Dellabona P., Allen P., Benoist C. and Mathis D. (1990) Delineation of antigen contact residues on an MHC class II molecule. *EMBO J.* **13**, 4215–4223.

Porcelli S., Morita C. T. and Brenner M. B. (1992) CD1β restricts the response of human CD4$^-$8$^-$ T lymphocytes to a microbial antigen. *Nature* **360**, 593–597.

Rao A., Faas S. J. and Cantor H. (1984a) Activation specificity of arsonate-reactive T cell clones. *J. Exp. Med.* **159**, 479–494.

Rao A., Ko W. W.-P., Faas S. J. and Cantor H. (1984b) Binding of antigen in the absence of histocompatibility proteins by arsonate-reactive T-cell clones. *Cell* **36**, 879–888.

Sambrook J., Fritsch E. F. and Manjatis T. (1989) *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.

Scott D. L., Otwinowski Z., Gelb M. H. and Sigler P. B. (1990) Crystal structure of bee-venom phospholipase A_2 in a complex with a transition-state analogue. *Science* **250**, 1563–1566.

Siliciano R. F., Hemesath T. J., Pratt J. C., Dintzis R. Z., Dintzes H. M., Acuto O., Shin H. S. and Reinherz E. L. (1986) Direct evidence for the existence of nominal antigen binding sites on T cell surface Ti α–β heterodimers of MHC-restricted T cell clones. *Cell* 47, 161–171.

Smith G. P. (1985) Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science* 228, 1315–1317.

White J., Herman A., Pullen A. M., Kubo R., Kappler J. W. and Marrack P. (1989) The V beta-specific superantigen staphylococcal enterotoxin B: stimulation of mature T cells and clonal deletion in neonatal mice. *Cell* 56, 27–35.

White J., Pullen A., Choi K., Marrack P. and Kappler J. W. (1993) Antigen recognition properties of mutant Vβ3$^+$ T cell receptors are consistent with an immunoglobulin-like structure for the receptor. *J. exp. Med.* 177, 119–125.

Winter G., Griffiths A. D., Hawkins R. E. and Hoogenboom H. R. (1994) Making antibodies by phage display technology. *A. Rev. Immun.* 12, 433–455.

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 12654–12658, December 1994
Immunology

Functional three-domain single-chain T-cell receptors

SHAN CHUNG†‡, KAI W. WUCHERPFENNIG†, STEVEN M. FRIEDMAN§, DAVID A. HAFLER¶, AND JACK L. STROMINGER†

†Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138; §Department of Medicine, Cornell University Medical College, New York, NY 10021; and ¶Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

Contributed by Jack L. Strominger, July 19, 1994

ABSTRACT **T-cell receptors (TCRs) are membrane anchored heterodimers structurally related to antibody molecules. Single-chain antibodies can be engineered by linking the two variable domains, which fold properly by themselves. However, proper assembly of the variable domains of a human TCR (V_α and V_β) that recognize the HLA-DR2b/myelin basic protein-(85–99) peptide complex was critically dependent on the addition of a third domain, the constant region of the TCR β chain (C_β), to the single-chain construct. Single-chain molecules with the three-domain design, but not those with the two-domain design, expressed in a eukaryotic cell as chimeric molecules linked either to glycosyl phosphatidylinositol or to the transmembrane/cytoplasmic domains of the CD3 ζ chain were recognized by a conformation-sensitive monoclonal antibody. The chimeric three-domain single-chain TCR linked to CD3 ζ chain signaled in response to both the specific HLA-DR/peptide and the HLA-DR/superantigen staphylococcal enterotoxin B complexes. Thus, by using this three-domain design, functional single-chain TCR molecules were expressed with high efficiency. The lipid-linked single-chain TCR was solubilized by enzymatic cleavage and purified by affinity chromatography. The apparent requirement of the constant domain for cooperative folding of the two TCR variable domains may reflect significant structural differences between TCR and antibody molecules.**

T-cell receptor (TCR) recognition of antigen fragments presented by major histocompatibility complex (MHC) molecules is a critical step in the initiation of a specific immune response (1, 2). The TCR α and β chains are each composed of two immunoglobulin-like domains; most of the amino acid residues that are found to be highly conserved in the variable (V) region of immunoglobulins are also found in TCR V regions, suggesting that the tertiary structure of the TCR may resemble that of immunoglobulins (3, 4). However, TCR V regions have significantly more primary sequence variability, an increased apparent rate of divergence in phylogeny, and peaks of variability in addition to those noted in immunoglobulins (5, 6). In order to understand and control the molecular interactions underlying T-cell recognition of MHC/peptide complexes, complete structural knowledge of the TCR is required.

Several approaches have been employed to produce soluble, recombinant TCRs. In these recombinant TCR molecules, the transmembrane/cytoplasmic regions of α and β chains were replaced with sequences from lipid-linked proteins (7), the CD3 ζ chain (8), or immunoglobulins (9, 10). Soluble TCRs were either recovered as secreted proteins or obtained by enzymatic cleavage of the surface-expressed recombinant proteins. All of these approaches rely, however, on the assembly of the heterodimer, which is inefficient (11). In addition, high-level expression of the human TCR α chain in transfected eukaryotic cells is not stable. These problems can be avoided by the design of a single-chain (sc) recombinant protein in which the V regions of the heterodimer are joined by a short peptide linker. Such a design has been successfully applied to antibody molecules (12). Such recombinant molecules, scFv, have a specificity and affinity similar to that of native antibodies (12). Several reports have described the production of scTCRs in bacterial expression systems using the sc antibody (Fv) design (13–16), but none have presented functional data indicating that these scTCRs could recognize their MHC/peptide complexes or superantigens. Recently, however, the production in bacteria of a scTCR with the two-domain Fv design that could recognize its natural ligands has been reported, although the fraction correctly refolded was extremely low (17).

In the present report, different scTCR designs were evaluated in transfected eukaryotic cells with respect to surface expression of TCR molecules, proper folding, and recognition of the appropriate MHC/peptide ligand. A three-domain sc construct [α-chain V (V_α)–linker–β-chain V (V_β)–β-chain constant (C_β)] was stably expressed on the cell surface when linked to a glycosyl phosphatidylinositol (GPI) anchor and recognized by a conformation-dependent monoclonal antibody (mAb) specific for the V_β17 segment. The soluble form of this recombinant protein could be readily obtained by enzymatic cleavage with phosphatidylinositol-specific phospholipase C (PI-PLC). Replacement of the GPI domain with the cytoplasmic portion of the ζ chain resulted in a functional TCR molecule that transduced an intracellular signal following recognition of either the proper MHC/peptide or the MHC/staphylococcal enterotoxin B (SEB) superantigen complex. The production of a functional scTCR directly demonstrates the feasibility of employing sc design to produce soluble TCRs.

MATERIALS AND METHODS

Construction of Recombinant TCR Molecules. cDNAs of TCR α and β chains were prepared from mRNA of Hy.2H9 cells (18) with Superscript reverse transcriptase (BRL) and an oligo(dT) primer (Sigma) and were amplified by PCR using Vent DNA polymerase (New England Biolabs) and primers 5′-GCTCGAGGCGGCGATGGAAACTCTCCTGGGAGT-3′ (A5) and 5′-GGAATTCAGCTGGACCACAGCCGC-3′ for α-chain and 5′-GCTCGAGCTCTGCCATGGACTCCTG-GA-3′ and 5′-GGAATTCAGAAATCCTTTCTCTTGAC-3′ for β-chain. The cDNAs were cloned into the mammalian expression vector pBJ-neo (8). GPI-anchored TCR molecules (α-PI and β-PI) were constructed as follows. A *Ban* I site was

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TCR, T-cell receptor; MHC, major histocompatibility complex; sc, single chain; GPI, glycosyl phosphatidylinositol; PI-PLC, phosphatidylinositol-specific phospholipase C; SEB, staphylococcal enterotoxin B; MBP, myelin basic protein; V, variable; C, constant; IL-2, interleukin 2; mAb, monoclonal antibody.

‡Present address: Cystar, Inc., 30 Commerce Way, Woburn, MA 01801.

engineered after the fifth amino acid residue beyond the last cysteine by oligonucleotide-directed mutagenesis. The region 3′ of the *Ban* I site was then replaced with a *Ban* I–*Not* I fragment encoding the GPI signal domain from the human placental alkaline phosphatase. For the construction of various scTCRs, variable domains of the 2H9 TCR α and β chains were prepared by PCR using primers A5 and 5′-CAGAGCTCACGGATGAACAATAAGGCTGGT-3′ for the V_α domain in all the scTCR constructs, 5′-TCGGATATCGATGGTGGAATCACTCAGTCC-3′ (B5) and 5′-CAGAGATCAGCACGGTGAGCCGGTTCCC-3′ for the V_β domain in AB-PI-1, 5′-GTGGGAGATCTCTGCTTCTGATGGCTCAAAC and B5 for the V_β domain in AB-PI-2, 5′-CACGGATCCCCGTCTGCTCTACCCCAGGC and B5 for the V_β and C_β domains in ABC-PI, and 5′-CACGGATCCCCGTCTGCTCTACCCCAGGC-3′ and B5 for the V_β and C_β domains in ABC-ζ. The cDNA encoding the transmembrane and cytoplasmic domains of murine CD3 ζ chain (8) was a gift of R. D. Klausner (National Institutes of Health). Convenient restriction sites were engineered at the end of each fragment to aid in the assembly of the construct. The linker was a 15-amino acid motif of GGGGS repeated three times (12) with *Sac* I at the 5′ end and *Eco*RV at the 3′ end. Except for α-PI, all the constructs were cloned into pBJ-neo, which carries the G418-resistance gene. α-PI was cloned into pCEP-4 (Invitrogen), which bears the hygromycin-resistance gene. All constructs were verified by multiple restriction digests and by sequencing with the Sequenase kit (United States Biochemical).

Affinity Purification and Characterization of a Soluble Three-Domain scTCR. After transfection and G418 selection (8), cells expressing a high level of GPI-linked three-domain scTCR (ABC-PI) were isolated by three rounds of sorting. The resulting cells were grown in spinner culture to a density of 10^6 per ml and harvested by centrifugation. The pellet was washed twice with phosphate-buffered saline (PBS) and resuspended in PBS containing 2 mM Pefabloc (Centerchem, Stamford, CT) to a density of 5×10^7 per ml with PI-PLC (Sigma) added at 1 unit/ml. Cells were incubated at 37°C for 1 hr with constant rocking. The supernatant was collected by centrifugation and by passage through a 0.45-μm filter and applied to a column of Acti-gel (Sterogen, Arcadia, CA) with immobilized βF1. The column was washed with 10 volumes of PBS and the soluble TCR was eluted with 0.15 M glycine (pH 2.8). Fractions were immediately neutralized with 0.1 volume of saturated Tris. The soluble TCR was then dialyzed against >100 volumes of PBS at 4°C with at least four changes and concentrated to 0.5 mg/ml by vacuum dialysis against PBS. Five micrograms of purified soluble three-domain scTCR was analyzed by SDS/PAGE under reducing conditions.

Stimulation of Transfectants with Antibodies, SEB, and Peptide/MHC Ligands. ABC-ζ-transfected BW5147β $\alpha^-\beta^-$ (19) cells (5×10^4 per well) were cultured in a 96-well round-bottom plate to which various antibodies had been immobilized (1 μg per well). The supernatants were collected after 24 hr and interleukin 2 (IL-2) production was assessed in a bioassay using an IL-2-dependent cell line (CTLL) and the CellTiter-96 nonradioactive proliferation assay (Promega). In the case of ABC-ζ-transfected RBL-2H3 (8) cells, the cells were incubated with [^{3}H]serotonin (NEN) at 0.5 μCi (18.5 kBq) for 24 hr before they were added to the antibody plate. After incubation at 37°C for 2 hr, radioactivity released into the supernatant was measured in a liquid scintillation counter. The specific serotonin release was calculated as described (8). For SEB stimulation, 5×10^4 transfected cells per well were cultured with various concentrations of SEB (Toxin Technology, Sarasota, FL) in the presence or absence of 2×10^5 B cells. For antigen presentation, 5×10^4 transfected cells per well were cocultured with 2×10^5 B cells which were incubated with or without the myelin basic protein (MBP)-(85–99) peptide for 3 hr before the experiment. The assays were conducted as described above.

RESULTS

mAb C1 Recognizes a Conformational Epitope of TCR. Recombinant TCR molecules were generated by using the TCR α- and β-chain sequences of the human MBP-specific T-cell clone Hy.2H9 (18). This clone TCR is composed of the V_α3.1 and V_β17.1 segments and is specific for the immunodominant MBP peptide MBP-(85–99) in the context of HLA-DR2 (DRA, DRB1*1602) (18). Usage of the V_β17.1 segment allowed the proper folding of recombinant TCRs to be probed with the superantigen SEB (20) and the mAb C1 (21). To confirm the V_β17 specificity of C1, the extracellular domains of TCR α and β chains of Hy.2H9 cells were fused to the C-terminal sequence from human placental alkaline phosphatase for GPI anchorage (Fig. 1) and the DNAs encoding the GPI-anchored β and α chains (β-PI and α-PI) were sequentially transfected by electroporation (8) into a TCR α- and β-chain-deficient murine lymphoma cell line, BW5147 $\alpha^-\beta^-$ (BW$^-$) (19). The surface expression of the GPI-anchored TCR chains was monitored by staining with mAbs αF1 (22), βF1 (23), and C1. αF1 and βF1 recognize nonconformational epitopes located in the C region of the TCR α and β chains, respectively. Surface expression of the GPI-anchored TCR β chain is independent of heterodimer formation and assembly of the CD3 complex (7). In the β-PI-transfected cells (Fig. 2, open curves), high-level expression of β-PI was confirmed by staining with βF1. Interestingly, there was little C1 staining of these transfectants. However, when a GPI-anchored 2H9 α chain was supertransfected into the β-PI transfectant (Fig. 2, shaded curves), C1 reactivity was greatly increased while the level of βF1 staining remained constant. Thus, the V_β17-specific C1 epitope is conformational and dependent on the proper pairing of TCR α and β chains and can therefore be used to assess the proper folding of recombinant TCRs bearing a V_β17 sequence.

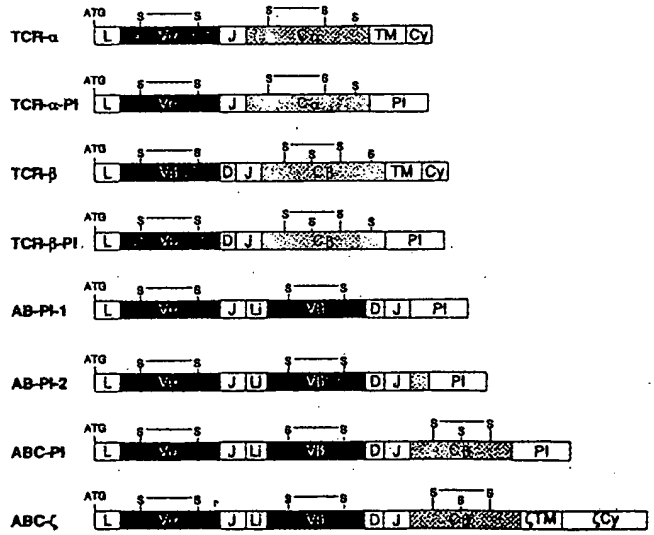


FIG. 1. Schematic representation of TCR α- and β-chain genes and various chimeric constructs. S—S, disulfide bond; L, leader; V, variable segment; J, joining segment; C, constant region; TM, transmembrane region; Cy, cytoplasmic region; ATG, start codon; Li, 15-residue peptide linker containing three repeats of GGGGS; PI, GPI domain of human placental alkaline phosphatase with the sequence LAPPAGTTDAAHPGRSVVPALLPLLAGTLLLL (7). The ζ region contains transmembrane and cytoplasmic domains of the murine CD3 ζ chain starting at position 31 (8).

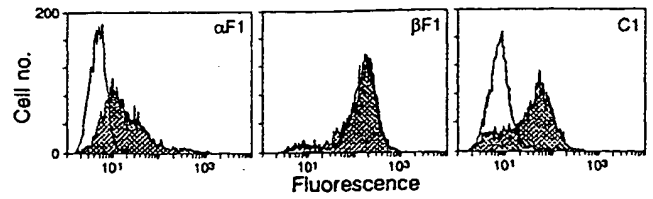


FIG. 2. Coexpression of α-PI and β-PI are required for C1 reactivity. Flow cytomeric analysis of β-PI-transfected BW⁻ cells stained with mAb αF1, βF1, or C1 before (open curves) and after (shaded curves) the supertransfection of the α-PI construct.

High-level expression of the TCR α chain (α-PI) was, however, not stable either alone or in the presence of β-PI. Attempts were made on several cell lines, including COS-7, CHO-K1, and a TCR-deficient variant of Jurkat cells, JK-β⁻ (J.RT3-T3.5, American Type Culture Collection). The expression level of α-PI was comparable to that of β-PI after the initial drug selection, but continued culture for less than a month yielded a population of cells with little surface expression of α-PI, whereas β-PI expression was stable (data not shown). The inability to obtain cell lines with stable high-level expression of the PI-anchored human TCR α-chain has been reported by other laboratories as well (24).

Expression and Purification of a Three-Domain scTCR. To overcome the limitations set by the unstable expression of the human TCR α chain, various sc designs were examined. Initially, a design similar to that of sc antibodies (Fv) was chosen (12). A 15-residue flexible linker was used to link the C terminus of the V_α–J_α domain to the N terminus of the β chain. The GPI domain was then ligated to the C terminus of the V_β–J_β domain. The construct (AB-PI-1, Fig. 1) was transfected into several cell lines, including JK-β⁻, COS-7, CHO-K1, and BW⁻. Although the expression of the gene was confirmed by the detection of the correct RNA transcripts (Fig. 3*B*), no surface expression was detected, as evidenced by negative C1 antibody staining (Fig. 3*A*). Immunoprecipitation after metabolic labeling failed to recover any C1-reactive sc molecules from these transfectants (data not shown). The inability to identify any C1-reactive protein could have been due to the design of this molecule, such as insufficient linker length between the extracellular domain and the GPI domain. To improve the accessibility of the sc construct, another two-domain scTCR was designed in which an extra 30-amino acid portion of the N terminus of the C_β domain was added as a hinge region. The transfectants of this construct (AB-PI-2, Fig. 1) were still not reactive with the C1 antibody (data not shown). Finally, the entire C_β domain was added to the sc construct. A complete C_β domain should provide enough distance for the V_α–V_β domains to be expressed on the cell surface and, more importantly, should allow surface expression to be monitored with another antibody, βF1 (23). This three-domain scTCR was constructed by extending the TCR β-chain sequences to the residue right before the last cysteine (the sixth cysteine), which was then fused to the GPI domain. The last cysteine was deleted to prevent dimerization between C_β domains. This construct (ABC-PI, Fig. 1) was transfected into BW⁻ cells and surface expression was confirmed by staining with both βF1 and C1 (Fig. 3*C*, shaded curves). Both antibodies stained the cells with comparable efficiency, suggesting that most of the molecules were expressed in the correct conformation. Moreover, the molecule could be efficiently cleaved from the cell surface with PI-PLC (Fig. 3*C*, open curves).

Soluble three-domain scTCR was purified from transfectants after PI-PLC cleavage followed by affinity chromatography using the βF1 antibody. The purified three-domain scTCR appeared as multiple bands at 50–70 kDa after SDS/PAGE (Fig. 3*D*). The heterogeneity of scTCR is probably the

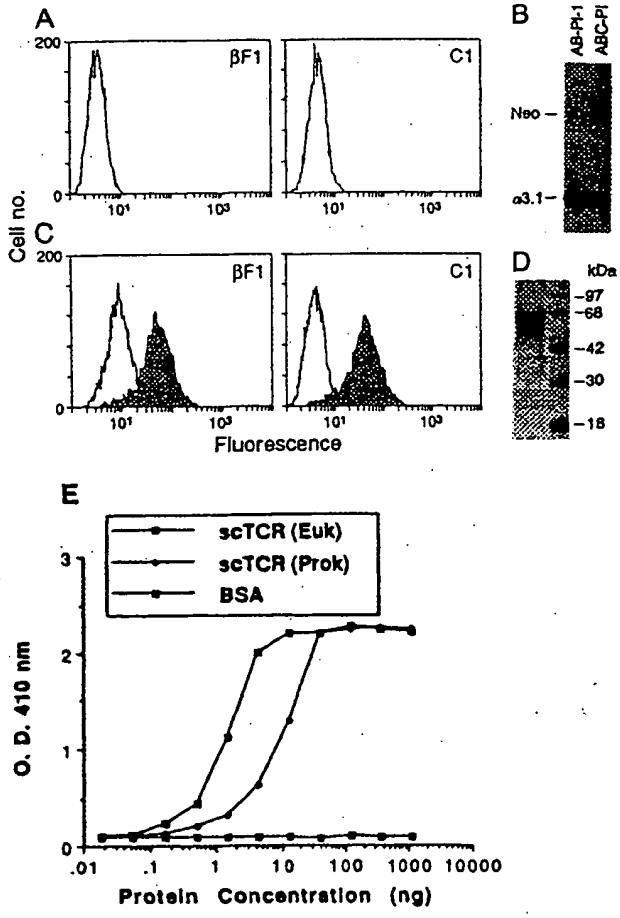


FIG. 3. Flow cytomeric analysis of surface expression of scTCR constructs. (*A*) Lack of C1 reactivity in BW⁻ cells transfected with two-domain scTCR (AB-PI-1). (*B*) RNA analysis of poly(A)-enriched total cell RNA isolated from BW⁻ cells transfected with two-domain and three-domain scTCR constructs AB-PI-1 and ABC-PI, respectively. Samples were analyzed in a nuclease S1 protection assay (25) using probes specific for the 5′ end of the transcripts from the TCR constructs. The coexpressed *neo* (G418-resistance gene) transcripts from the vector were analyzed with a probe at the same time as a control. (*C*) Flow cytometric analysis of ABC-PI-transfected BW⁻ cells with both βF1 and C1 antibodies before (shaded curves) and after (open curves) PI-PLC treatment. (*D*) SDS/PAGE of affinity-purified three-domain scTCR. (*E*) Comparison of C1 reactivity of three-domain scTCRs produced from eukaryotic (Euk) and prokaryotic (Prok) expression systems in a two-antibody ELISA. A plateau is reached because the amount of βF1 attached to the plate became limiting.

result of variable glycosylation; its polypeptide size calculated from amino acid composition is 40 kDa. The structural integrity of the three-domain scTCR was verified by a two-antibody ELISA (Fig. 3*E*). The molecules were first captured by the βF1 antibody immobilized to the plate and then assessed for reactivity with the C1 antibody. When compared with the three-domain scTCR produced in a bacterial expression system (unpublished work), the scTCR from the eukaryotic system gave 10–20 times higher C1 reactivity. The purified three-domain scTCR was stable and could be stored in PBS at 4°C for months without significant loss of C1 reactivity.

Functional Characterization of a Chimeric Three-Domain scTCR. To directly assess the functional integrity of the three-domain scTCR, a self-signaling scTCR was produced by replacing the GPI domain with the transmembrane and cytoplasmic domains of the CD3 ζ chain. These regions have been shown to be sufficient for signal transduction when its

extracellular fusion partner is crosslinked by an antibody or by the proper ligand (8, 26, 27). To enable the recovery of three-domain scTCR as a soluble form, a linker containing a thrombin cleavage site was inserted into the junction of three-domain scTCR and the ζ domain. The construct (ABC-ζ) was transfected into BW$^-$ cells (28) and the rat basophilic leukemia cell line RBL-2H3 (RBL) (8), and the populations displaying high-level expression of three-domain scTCR were isolated by three rounds of cytofluorometric sorting using the antibody βF1. The ABC-ζ-transfected cells were first stimulated with various antibodies to confirm the self-signaling nature of this recombinant molecule. The signal transduced upon the activation of the three-domain scTCR was measured as IL-2 production in BW$^-$ transfectants, whereas serotonin release was measured in RBL transfectants. Both transfectants showed a strong response following βF1 and C1 stimulation but not to purified mouse immunoglobulin or anti-CD8 antibody used as controls (Fig. 4 *A* and *B*). The structural integrity of the scTCR was further examined with the superantigen SEB, which binds to both V$_\beta$17 and MHC class II molecules, resulting in TCR crosslinking and T-cell activation regardless of the peptide bound to the MHC molecule (20, 29). ABC-ζ transfectants displayed a concentration-dependent response toward SEB (Fig. 4*C*) when the superantigen was presented by transformed B-cell lines with high-level expression of DR1 (DRA, DRB1*0101; cell line LG2) or DR2 (DRA, DRB1*1602; cell line 9016). Thus, the lateral face of the TCR V$_\beta$ region to which SEB is thought to bind (30) is structurally intact.

To prove that the three-domain scTCR did indeed recognize the MHC/peptide ligand, antigen presentation experiments using the natural ligand for the Hy.2H9 clone, 9016 cells bearing the DRB1*1602 allele of DR2, and MBP-(85–99) peptide were performed. To ensure detection of subtle abnormalities in the structure of the three-domain scTCR, 9009 cells (DRA, DRB1*1601), which also bind MBP-(85–99), were used as a control. DRB1*1601 and DRB1*1602 differ only at position 67 in the DR$_\beta$1 domain; this TCR contact-residue substitution does, however, abolish recognition of the peptide by the parent T-cell clone (ref. 18; K.W.W.,

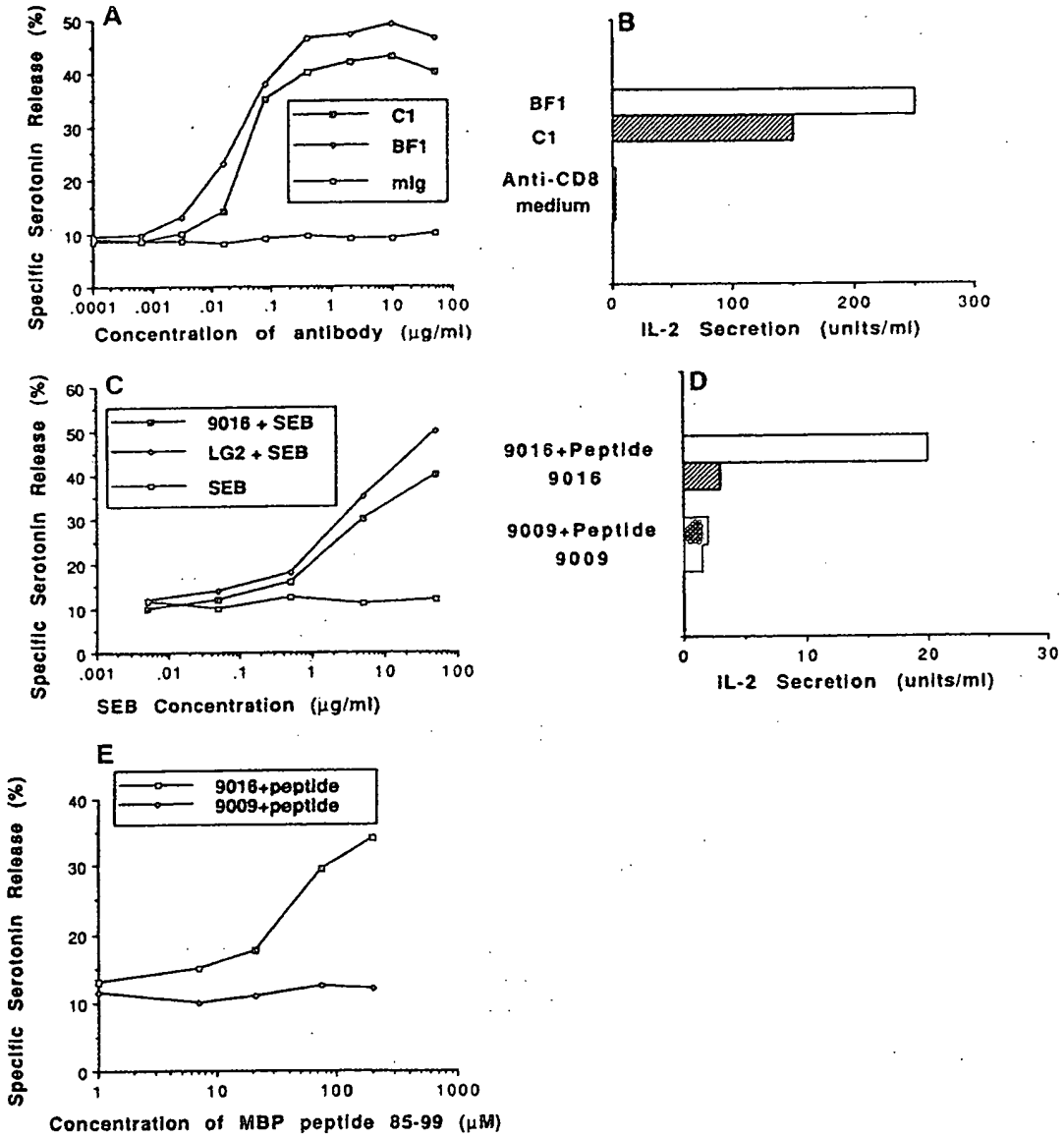


FIG. 4. (*A*) [^{3}H]Serotonin release following TCR crosslinking of three-domain scTCR–ζ chimeric construct (ABC-ζ)-transfected RBL cells. C1 is specific to V$_\beta$17 and βF1 is specific to the C region of TCR β chain. mIg, mouse immunoglobulin. (*B*) IL-2 release following TCR crosslinking of ABC-ζ-transfected BW$^-$ cells. (*C*) [^{3}H]Serotonin release following SEB stimulation of ABC-ζ-transfected RBL cells. (*D*) MHC-restricted antigen-specific response from the ABC-ζ-transfected BW$^-$ cells. (*E*) Dose–response curve of MHC/peptide recognition of ABC-ζ-transfected RBL cells.

unpublished work). The ABC-ζ-transfected BW$^-$ cells secreted IL-2 in response to peptide-pulsed 9016 cells, but not to peptide-pulsed 9009 cells (Fig. 4*D*). Similar results were obtained with RBL transfectants (Fig. 4*E*), as serotonin release was dependent on the concentration of the MBP peptide used to pulse 9016 cells. The signal appeared to be weak when compared with antibody stimulation. This is not surprising, however, since saturating amounts of antibodies are expected to crosslink the majority of TCR molecules on the target cell, whereas a much smaller fraction of TCR molecules is probably engaged when T cells are cocultured with peptide-pulsed antigen-presenting cells, in which a maximum of 5–15% of the DR molecules bind the peptide. The requirement for a high concentration of peptide or SEB is not due to the sc design, since high concentrations of peptide are also needed to stimulate $\alpha\beta$ heterodimers of the TCR-ζ constructs (8). It is likely that the decrease in sensitivity results from the lack of CD3, CD4, and/or other adhesion/signaling molecules. Nonetheless, these results demonstrate that the three-domain scTCR was correctly folded and functionally competent. In addition, a soluble form of three-domain scTCR could be obtained from the ABC-ζ transfectants by thrombin cleavage and affinity purification (data not shown).

DISCUSSION

A scTCR molecule was designed which contains the V domains of both α and β chains and the C domain of the β chain. This scTCR molecule could be stably expressed at a high level in eukaryotic cells and could be isolated in a soluble form by enzymatic cleavage and affinity chromatography. The V_α and V_β domains appeared to be properly paired, since the scTCR bound to a conformation-dependent mAb, the superantigen SEB, and the proper MHC/peptide ligand. This design of scTCR offers an alternative to the two-chain design of soluble TCRs and has several advantages. (*i*) The sc design avoids the low-efficiency dimerization process which may be the limiting step in the assembly of TCR heterodimers from α and β subunits made in *Escherichia coli*. The sc design therefore allows efficient expression of the recombinant protein in quantities suitable for structural analysis and for some diagnostic or therapeutic applications. (*ii*) The design avoids the problems associated with the unstable expression of the human TCR α chain that have hindered efficient expression of human TCR molecules in eukaryotic cells. (*iii*) The sc design may allow the construction of TCR phage display libraries similar to those made for sc antibodies (28, 31). scTCR phage libraries may be powerful tools for the isolation of TCRs with defined specificities and/or high affinity for selective targeting of malignant and virally infected cells and for analyzing the interactions among TCR, MHC/peptide complexes, and superantigens.

Unlike antibodies, separately expressed V domains of TCR α and β chains have not been reconstituted to form heterodimers (ref. 32; K. L. Hilyard, personal communication). However, despite this success in producing the three-domain scTCR, a two-domain scTCR with detectable C1 reactivity could not be produced either in eukaryotic cells or in bacteria. The presence of the C_β domain (or part of the domain) may be required for the proper folding and/or stabilization of a scTCR molecule. In any event, the three-domain design provides a general means for the efficient production of functional scTCR molecules.

We thank Drs. James Huston and Herman Oppermann for helpful advice and Dr. Basya Rybalov for technical assistance. This work was supported by National Institutes of Health Grant CA47554 (J.L.S.), and National Multiple Sclerosis Society Postdoctoral Fellowship Award FG915-A-1 (S.C.). K.W.W. is a Harry Weaver Neuroscience Scholar of the National Multiple Sclerosis Society.

1. Ashwell, J. D. (1990) *Annu. Rev. Immunol.* **8**, 139–167.
2. Jorgensen, J. L., Reay, P. A., Ehrich, E. W. & Davis, M. M. (1992) *Annu. Rev. Immunol.* **10**, 835–837.
3. Davis, M. M. & Bjorkman, P. J. (1988) *Nature (London)* **334**, 395–402.
4. Chothia, C., Boswell, D. R. & Lesk, A. M. (1988) *EMBO J.* **7**, 3745–3755.
5. Behlke, M. A., Spinella, D. G., Chou, H. S., Sha, W., Hartl, D. L. & Loh, D. Y. (1985) *Science* **229**, 566–570.
6. Becker, D. M., Patten, P., Chien, Y.-H., Yokota, T., Eshaar, Z., Giedlin, M., Gascoigne, N. R. J., Goodnow, C., Wolf, R., Arai, A. & Davis, M. M. (1985) *Nature (London)* **317**, 430–434.
7. Lin, A. Y., Devaux, B., Green, A., Sagerstrom, C., Elliott, J. F. & Davis, M. M. (1990) *Science* **249**, 677–679.
8. Engel, I., Ottenhoff, T. H. M. & Klausner, R. D. (1992) *Science* **256**, 1318–1321.
9. Weber, S., Traunecker, A., Oloveri, F., Gerhard, W. & Karjalainen, K. (1992) *Nature (London)* **356**, 793–796.
10. Gregoire, C., Rebai, N., Schweisguth, F., Necker, A., Mazza, G., Auphan, N., Millward, A., Schmitt-Verhulst, A.-M. & Malissen, B. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8077–8081.
11. Klausner, R. D., Lippincott-Schwartz, J. & Bonifacino, J. S. (1990) *Annu. Rev. Cell Biol.* **6**, 403–431.
12. Huston, J., McCartney, J., Tai, M.-S., Mottola-Hartshorn, C., Jin, D., Warren, F., Keck, P. & Oppermann, H. (1992) *Int. Rev. Immunol.* **10**, 195–217.
13. Novotny, J., Ganju, R. K., Smiley, S. T., Hussey, R. E., Luther, M. A., Recny, M. A., Siliciano, R. F. & Reinherz, E. L. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 8646–8650.
14. Ward, E. S. (1992) *J. Mol. Biol.* **224**, 885–890.
15. Soo Hoo, W. F., Lacy, M. J., Denzin, L. K., Voss, E. W., Jr., Hardman, K. D. & Kranz, D. M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4759–4763.
16. Kurucz, I., Jost, C. R., George, A. J. T., Andrew, S. M. & Segal, D. M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 3830–3834.
17. Hilyard, K. L., Reyburn, H., Chung, S., Garboczi, D. N., Bell, J. I. & Strominger, J. L. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9057–9061.
18. Wucherpfennig, K. W., Zhang, J., Witek, C., Matsui, M., Modabber, Y., Ota, K. & Hafler, D. A. (1994) *J. Immunol.* **152**, 5581–5592.
19. White, J., Blackman, M., Bill, J., Kappler, J., Marrack, P., Gold, D. & Born, W. (1989) *J. Immunol.* **143**, 1822–1828.
20. Choi, Y., Kotzin, B., Herron, L., Callahan, J., Marrack, P. & Kappler, J. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8941–8945.
21. Friedman, S. M., Crow, M. K., Tumang, J. R., Tumang, M., Xu, Y., Hodtsev, A. S., Cole, B. C. & Posnett, D. N. (1991) *J. Exp. Med.* **174**, 891–900.
22. Henry, L., Tian, W. T., Rittershaus, C., Ko, J. L., Marsh, H. C., Jr., & Ip, S. H. (1990) *Hybridoma* **8**, 577–588.
23. Brenner, M. B., Mclean, J., Scheft, H., Warnke, R. A., Jones, N. & Strominger, J. L. (1987) *J. Immunol.* **138**, 1502–1509.
24. Devaux, B., Bjorkman, P. J., Stevenson, C., Greil, W., Elliott, J. F., Sagerstrom, C., Clayberger, C., Krensky, A. M. & Davis, M. M. (1991) *Eur. J. Immunol.* **21**, 2111–2119.
25. Chung, S. & Perry, R. P. (1993) *Nucleic Acids Res.* **21**, 3301–3308.
26. Irving, B. A. & Weiss, A. (1991) *Cell* **64**, 891–898.
27. Romeo, C. & Seed, B. (1991) *Cell* **64**, 1037–1045.
28. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991) *Science* **254**, 624–628.
29. Irwin, M. J., Hudson, K. P., Ames, K. T., Fraser, J. D. & Gascoigne, N. R. (1993) *Immunol. Rev.* **131**, 61–78.
30. Choi, Y., Herman, A., Digiusto, D., Wade, T., Marrack, P. & Kappler, J. (1990) *Nature (London)* **346**, 471–473.
31. McCafferty, J., Griffiths, A. D., Winter, G. & Chiswell, D. J. (1990) *Nature (London)* **348**, 552–554.
32. Ward, E. S. (1991) *Scand. J. Immunol.* **34**, 215–220.

US005759817A

United States Patent [19]

Barbas

[11] Patent Number: 5,759,817

[45] Date of Patent: Jun. 2, 1998

[54] **HETERODIMERIC RECEPTOR LIBRARIES USING PHAGEMIDS**

[75] Inventor: **Carlos Barbas**, San Diego, Calif.

[73] Assignee: **The Scripps Research Institute**, La Jolla, Calif.

[21] Appl. No.: **322,730**

[22] Filed: **Oct. 12, 1994**

Related U.S. Application Data

[63] Continuation of Ser. No. 826,623, Jan. 27, 1992, abandoned, which is a continuation-in-part of Ser. No. 683,602, Apr. 10, 1991, abandoned.

[51] **Int. Cl.**[6] **G01N 33/53**; C12N 15/00; C12P 21/00; C07K 16/00

[52] **U.S. Cl.** **435/69.7**; 435/7.1; 435/172.3; 435/320.1; 530/387.1; 530/387.3

[58] **Field of Search** ..... 69/1 T; 435/172.2, 435/5, 7.1, 172.1, 172.3, 320.1, 69.7; 536/24; 935/8; 530/387.1, 387.3, 412

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,223,409	6/1993	Ladner	435/69.7
5,403,484	4/1995	Ladner	435/235.1
5,427,908	6/1995	Dower et al.	435/5

FOREIGN PATENT DOCUMENTS

0 368 684	5/1990	European Pat. Off. .
WO91/17271	11/1991	WIPO .
WO91/18980	12/1991	WIPO .
WO91/19818	12/1991	WIPO .
WO92/01047	1/1992	WIPO .

OTHER PUBLICATIONS

Gussow, et al., *Cold Spring Harbor Symposia on Quantitative Biology*, vol. LIV, pp. 265–272 (1989).

Bass et al., *Proteins: Structure, Function and Genetics*, 8:309–314 (1990).

Bird et al., *Science*, 242:423–426 (1988).

Boss et al., *Nucl. Acids Res.*, 12:3791–3806 (1984).

Cabilly et al., *Proc. Natl. Acad. Sci.*, USA, 81:3273–3277 (1984).

Caton et al., *Proc. Natl. Acad. Sci., USA*, 87:6450–6454 (1990).

Cwirla et al., *Proc. Natl. Acad. Sci., USA*, 87:6378–6382 (1990).

Il'ichev et al., *Molekulyarnaya Biologiya*, 24:530–535 (1990).

Larrick et al., *Bio/Technology*, 7:934–938 (1989).

Model et al., "The Bacteriophages", ed: R. Calender, Plenum Publishing Corp., 2:375–456 (1988).

Mullinax et al., *Proc. Natl. Acad. Sci., USA*, 87:8095–8099 (1990).

Orlandi et al., Proc. Natl. Acad. Sci., USA, 86:3833–3837 (1989).

Parmely et al., Gene, 73:305–318 (1988).

Pluckthun, *Nature*, 347:497–498 (1990).

Rasched et al., *Microbiol. Rev.*, 50:401–427 (1986).

Sastry et al., *Proc. Natl. Acad. Sci.*, USA, 86:5728–5732 (1989).

Scott et al., *Science*, 249:386–390 (1990).

Short et al., Nuc. Acids Res., 16:7583–7600 (1988).

Skerra et al., *Science*, 240:1038–1041 (1988).

Skerra et al., *Bio/Technology*, 9:273–278 (1991).

Smith, G., *Science*, 228:1315–1317 (1985).

Takeda et al., *Nature*, 314:452–454 (1985).

Ward et al., Nature, 341:544–546 (1989).

Barbas et al., *Methods: A Companion to Methods in Enzymology*, 2:119–124 (1991).

Huse et al., *Science*, 246:1275–1281 (1989).

Kang et al., Proc. Natl. Acad. Sci., USA, 88:4363–4366 (1991).

McCafferty et al., *Nature*, 348:552–554 (1990).

Persson et al., *Proc. Natl. Acad. Sci., USA*, 88:2432–2436 (1991).

Barbas, C.F. and Lerner, R.A., "Combinational Immunoglobulin Libraries on the Surface of Phage (Phabs): Rapid Selection of Antigen-Specific Fabs" *Methods: A Companion to Methods in Enzymology*, 2:119–124 (1991).

Barbas, et al., "Assembly of Combinational Antibody Libraries on Phage Surfaces: The Gene III Site", *Proc. Natl. Acad. Sci. USA*, 88:7978–7982 (1991).

Kang, et al., "Linkage of Recognition and Replication Functions by Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces", *Proc. Natl. Acad. Sci. USA*, 88:4363–4366 (1991).

Higuchi 1989, in PCR Technology, Henry A. Erlich ed "M" Stockton Press.

Sambrook et. al. 1989 Molecular Cloning, pp. 15.51–15.84, Cold Springs Harbor Press, Cold Spring Harbor NY.

Primary Examiner—Nancy Degen
Assistant Examiner—Sean M. Garry
Attorney, Agent, or Firm—Thomas Fitting; Emily Holmes

[57] **ABSTRACT**

Filamentous phage comprising a matrix of cpVIII proteins encapsulating a genome encoding first and second polypeptides of an antogenously assembling receptor, such as an antibody, and a receptor comprised of the first and second polypeptides surface-integrated into the matrix via a cpVIII membrane anchor domain fused to at least one of the polypeptides with a mutagenized CDR3 region.

26 Claims, 12 Drawing Sheets

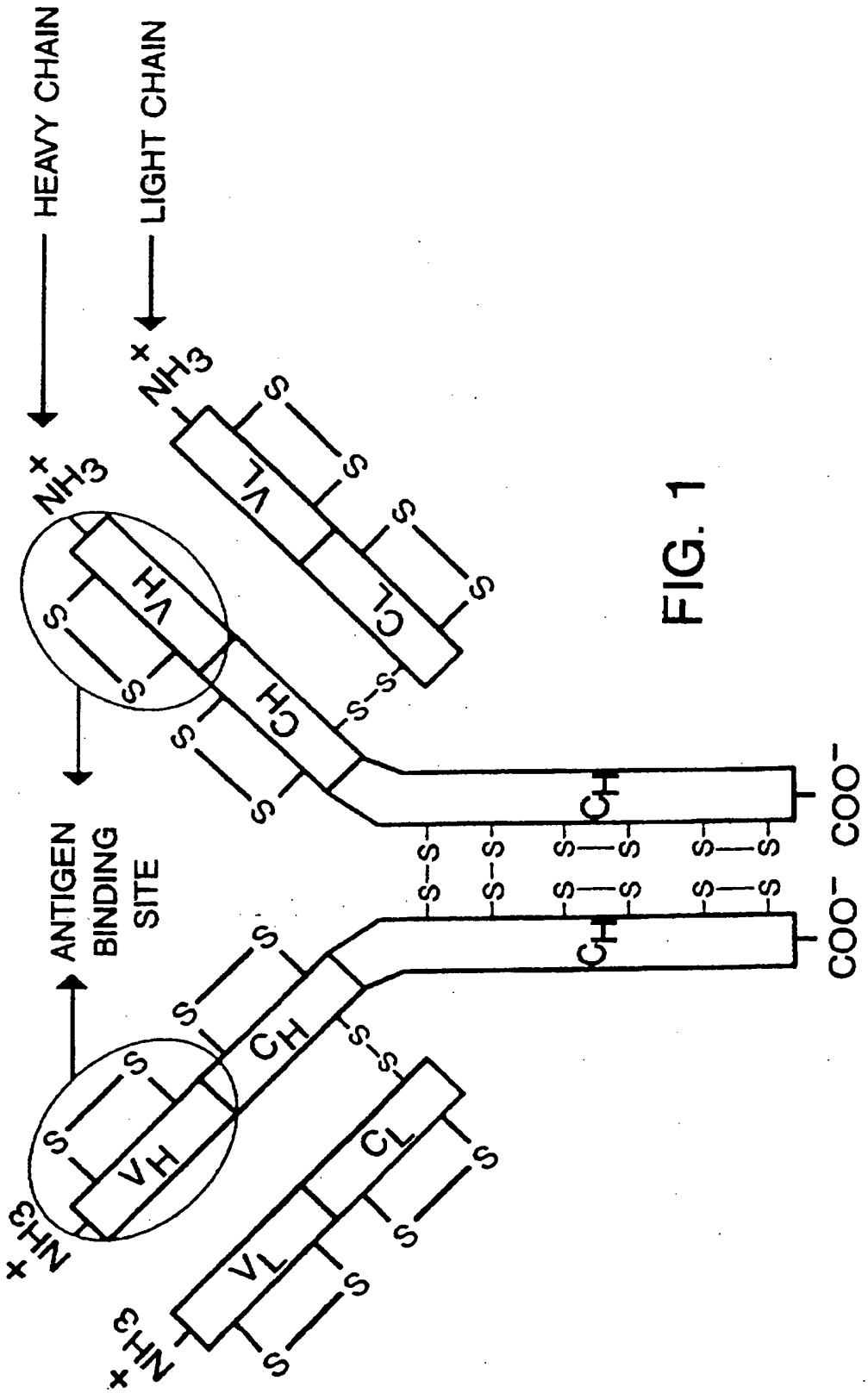

HEAVY CHAIN
LIGHT CHAIN
NH3
VL
CL
VH
CH
S
S-S
ANTIGEN BINDING SITE
COO-

FIG. 1

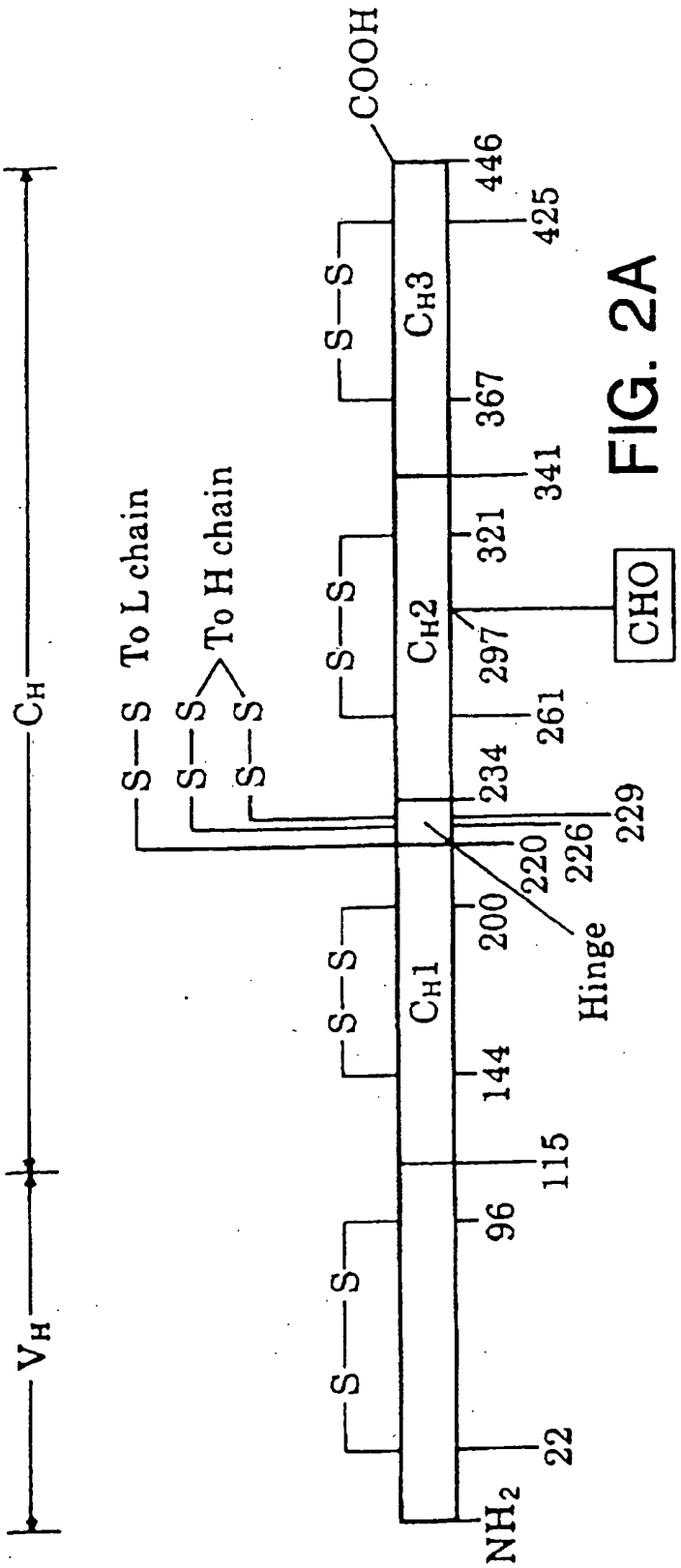

COOH
446
425
C_H3
367
341
321
C_H2
CHO
297
261
234
229
226
220
200
Hinge
C_H1
144
115
96
22
NH_2
S—S
To L chain
To H chain
C_H
V_H

FIG. 2A

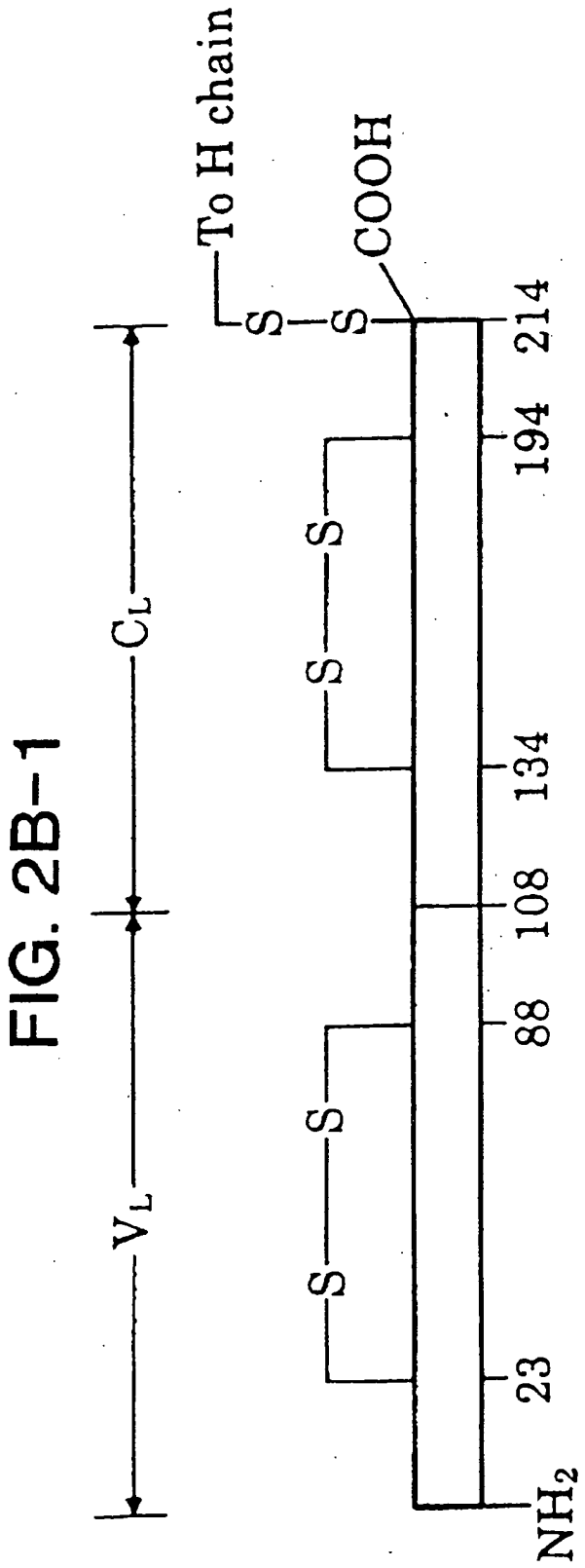


FIG. 2B-1

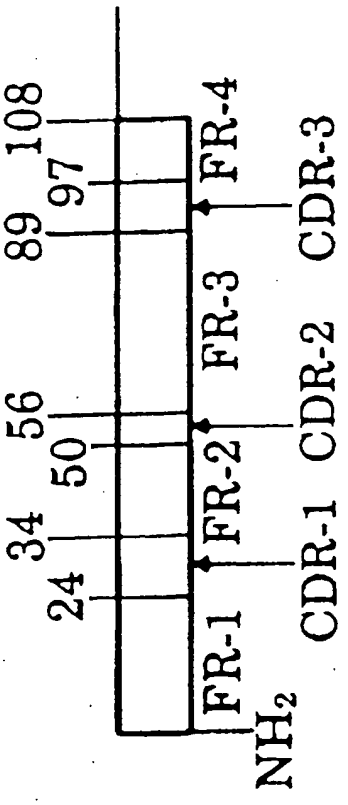


FIG. 2B-2

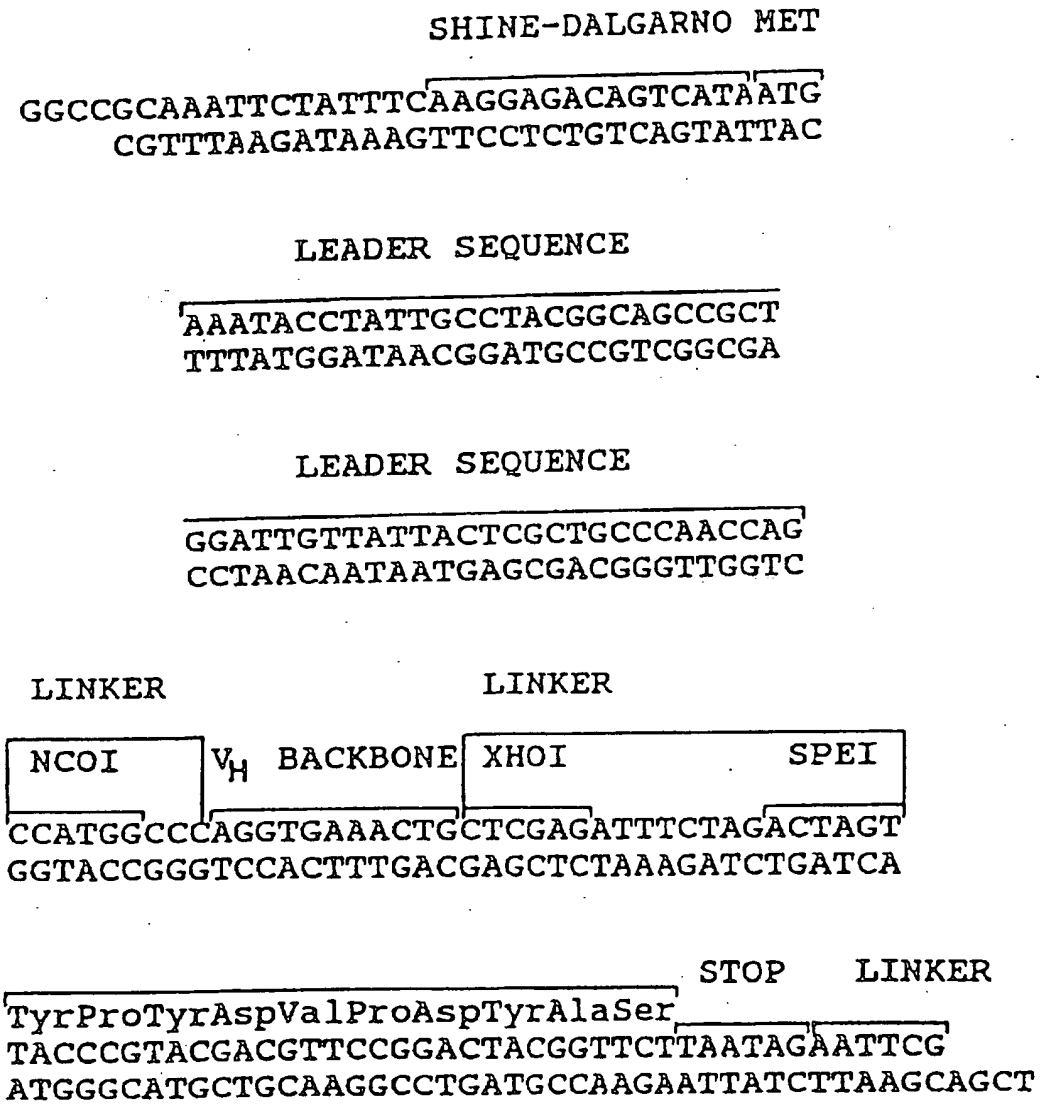

SHINE-DALGARNO MET
GGCCGCAAATTCTATTTCAAGGAGACAGTCATAATG
CGTTTAAGATAAAGTTCCTCTGTCAGTATTAC
LEADER SEQUENCE
AAATACCTATTGCCTACGGCAGCCGCT
TTTATGGATAACGGATGCCGTCGGCGA
LEADER SEQUENCE
GGATTGTTATTACTCGCTGCCCAACCAG
CCTAACAATAATGAGCGACGGGTTGGTC
LINKER
LINKER
NCOI
V_H BACKBONE
XHOI
SPEI
CCATGGCCCAGGTGAAACTGCTCGAGATTTCTAGACTAGT
GGTACCGGGTCCACTTTGACGAGCTCTAAAGATCTGATCA
STOP
LINKER
TyrProTyrAspValProAspTyrAlaSer
TACCCGTACGACGTTCCGGACTACGGTTCTTAATAGAATTCG
ATGGGCATGCTGCAAGGCCTGATGCCAAGAATTATCTTAAGCAGCT

FIG. 3

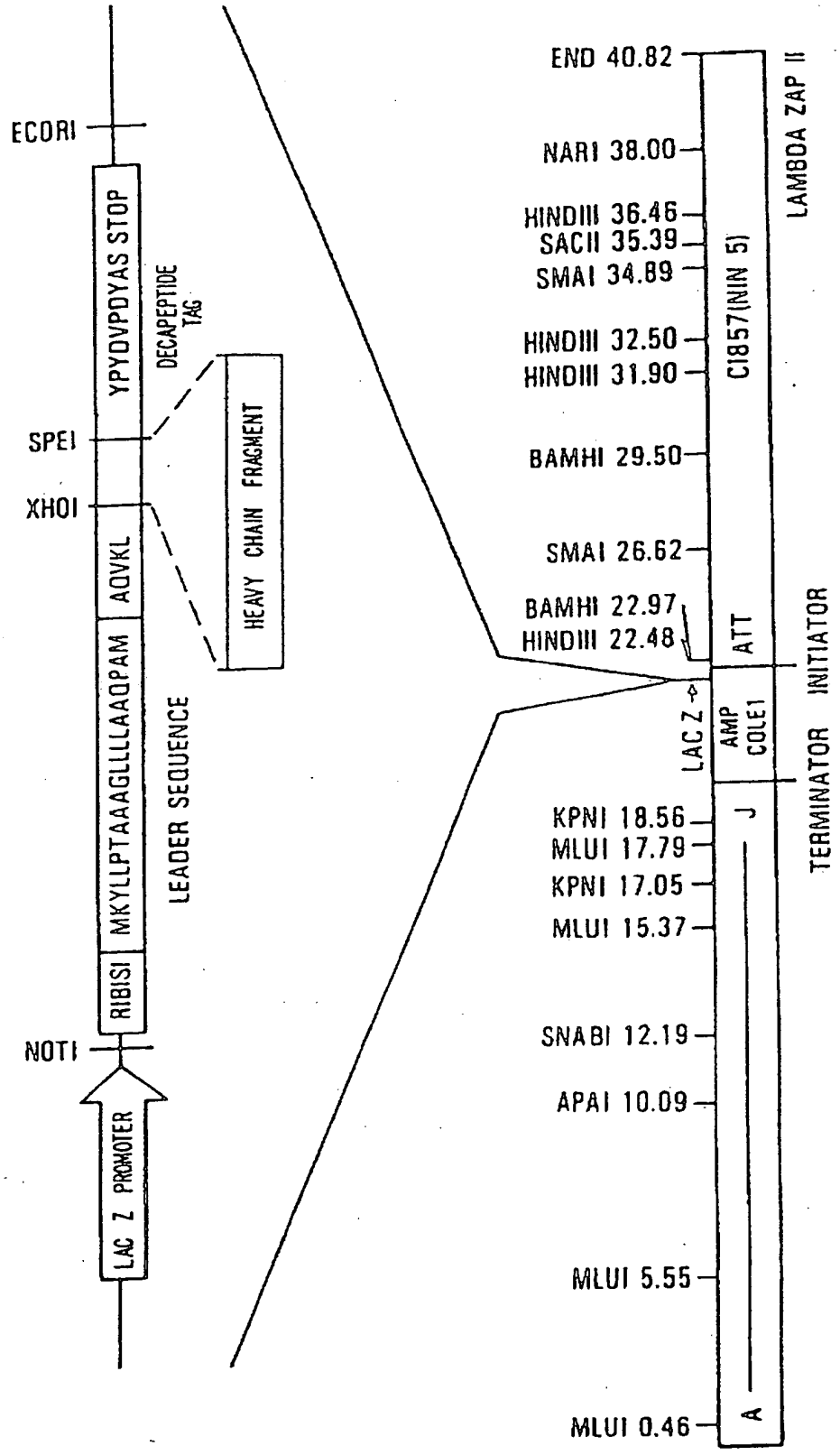

ECORI
YPYDVPDYAS STOP
DECAPEPTIDE TAG
SPEI
XHOI
AQVKL
HEAVY CHAIN FRAGMENT
MKYLLPTAAAGLLLAAQPAM
LEADER SEQUENCE
RIBISI
NOTI
LAC Z PROMOTER
END 40.82
NARI 38.00
HINDIII 36.46
SACII 35.39
SMAI 34.89
HINDIII 32.50
HINDIII 31.90
BAMHI 29.50
SMAI 26.62
BAMHI 22.97
HINDIII 22.48
C1857(NIN 5)
LAMBDA ZAP II
ATT
LAC Z
AMP
COLE1
TERMINATOR
INITIATOR
KPNI 18.56
MLUI 17.79
KPNI 17.05
MLUI 15.37
J
SNABI 12.19
APAI 10.09
MLUI 5.55
A
MLUI 0.46

FIG. 4

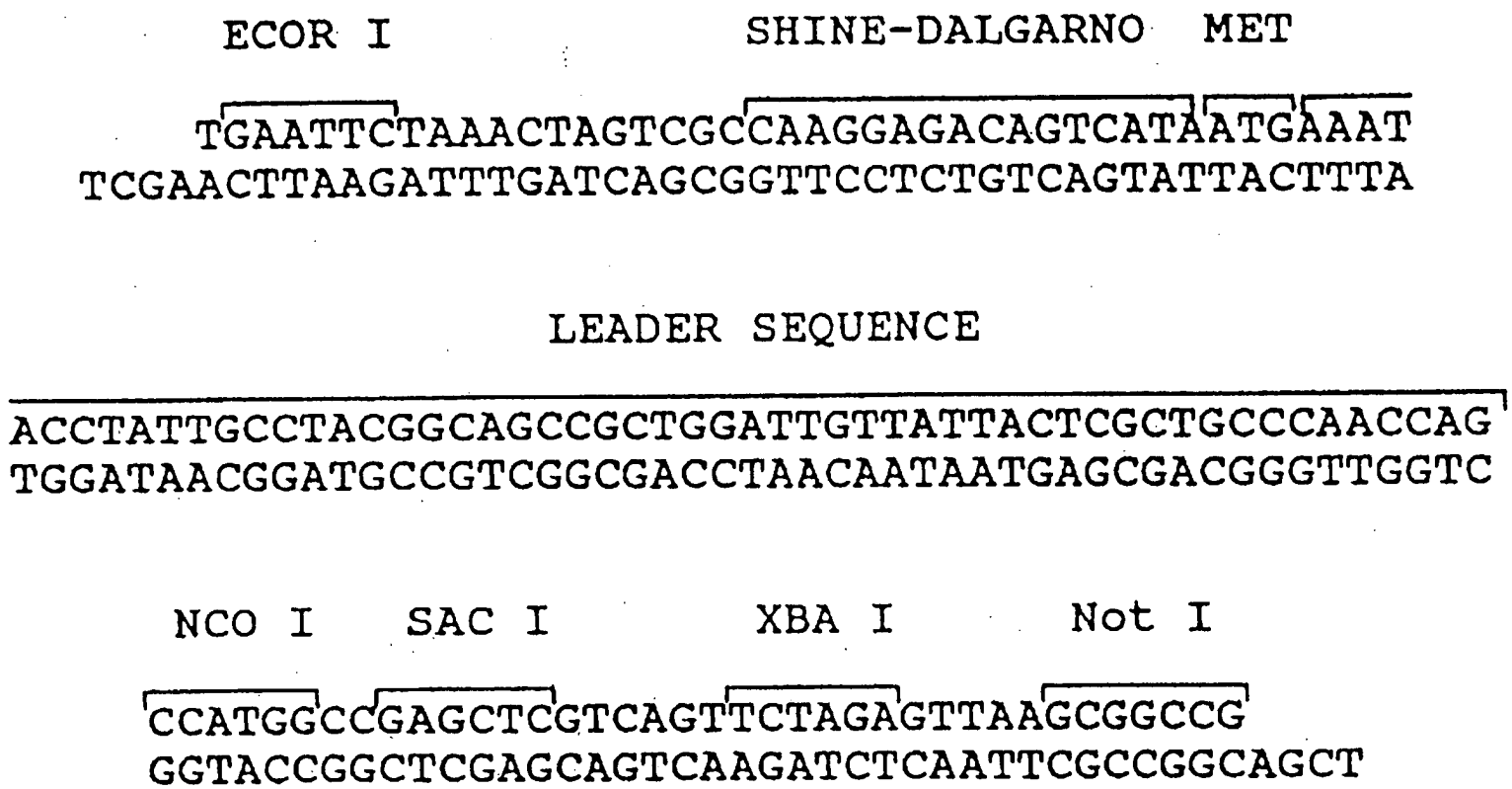

ECOR I
SHINE-DALGARNO
MET
TGAATTCTAAACTAGTCGCCAAGGAGACAGTCATAATGAAAT
TCGAACTTAAGATTTGATCAGCGGTTCCTCTGTCAGTATTACTTTA
LEADER SEQUENCE
ACCTATTGCCTACGGCAGCCGCTGGATTGTTATTACTCGCTGCCCAACCAG
TGGATAACGGATGCCGTCGGCGACCTAACAATAATGAGCGACGGGTTGGTC
NCO I
SAC I
XBA I
Not I
CCATGGCCGAGCTCGTCAGTTCTAGAGTTAAGCGGCCG
GGTACCGGCTCGAGCAGTCAAGATCTCAATTCGCCGGCAGCT

FIG. 5

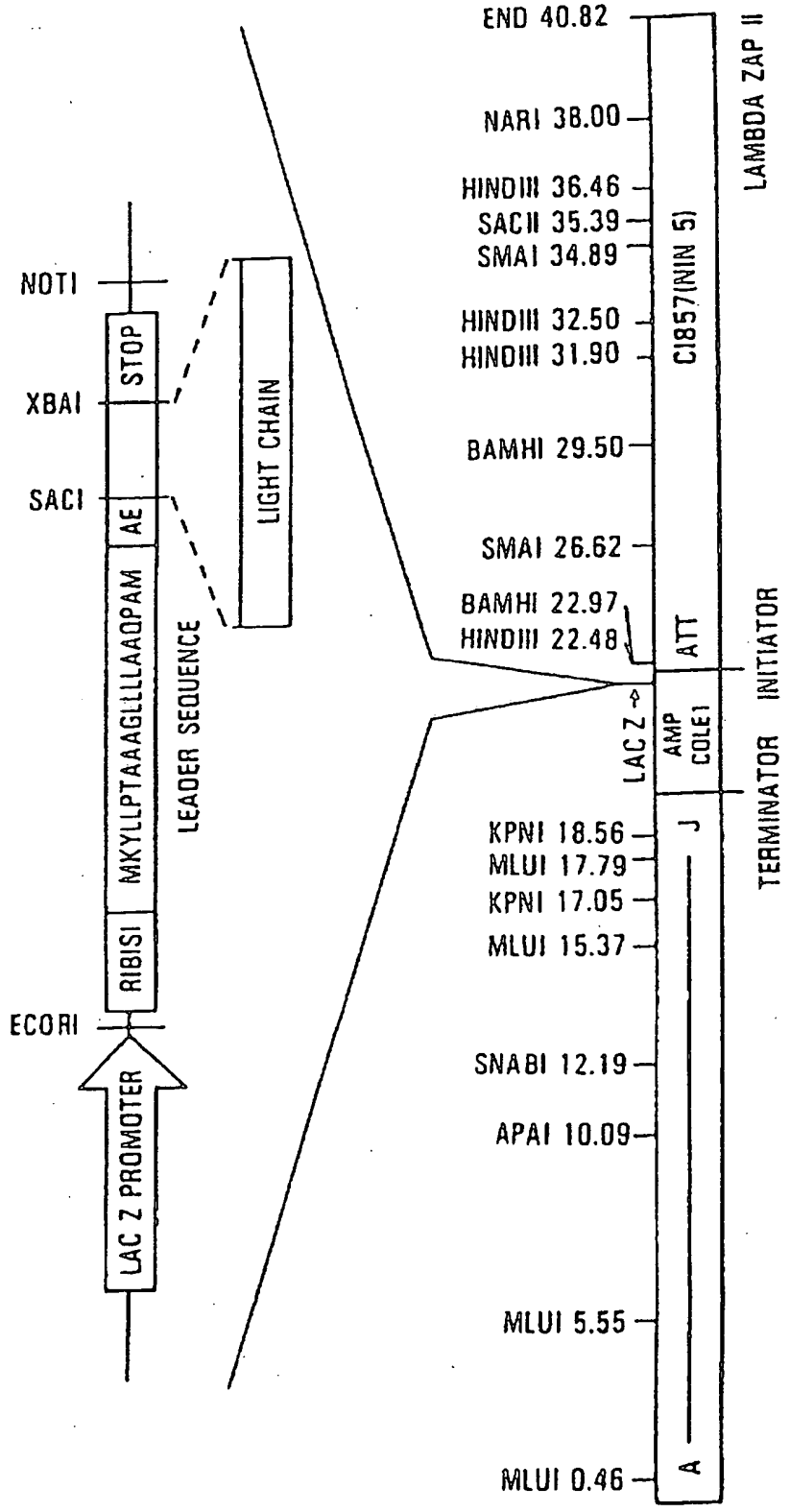

END 40.82
NARI 38.00
HINDIII 36.46
SACII 35.39
SMAI 34.89
HINDIII 32.50
HINDIII 31.90
BAMHI 29.50
SMAI 26.62
BAMHI 22.97
HINDIII 22.48
LAC Z
AMP
COLE1
KPNI 18.56
MLUI 17.79
KPNI 17.05
MLUI 15.37
SNABI 12.19
APAI 10.09
MLUI 5.55
MLUI 0.46
LAMBDA ZAP II
CI857(NIN 5)
ATT
TERMINATOR
INITIATOR
J
A
NOTI
XBAI
SACI
ECORI
STOP
AE
RIBISI
MKYLLPTAAAGLLLAAQPAM
LEADER SEQUENCE
LIGHT CHAIN
LAC Z PROMOTER

FIG. 6

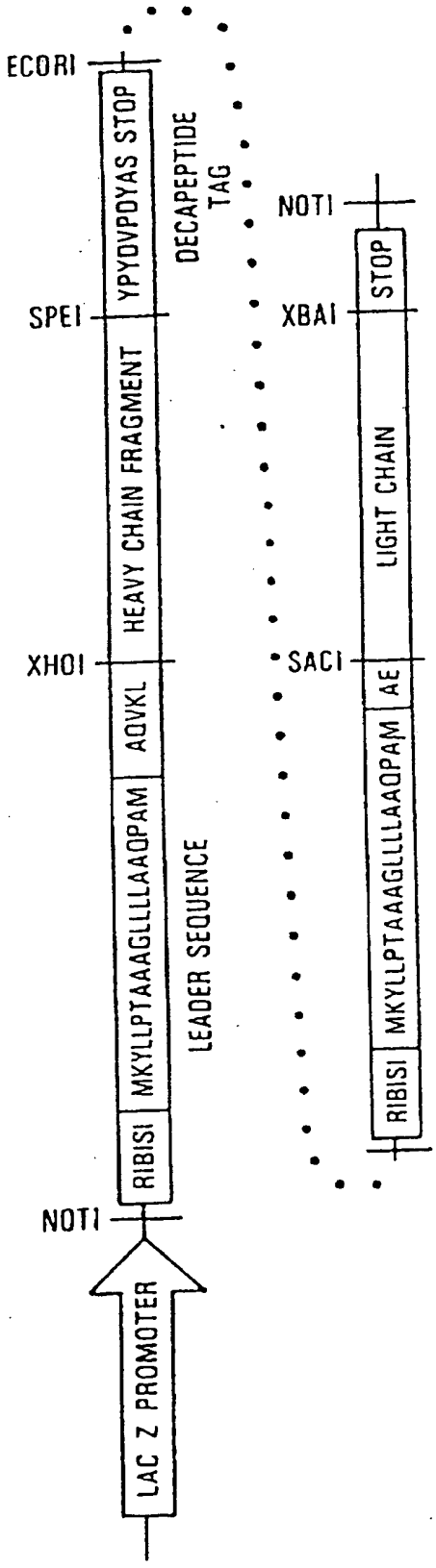

ECORI
SPEI
XHOI
NOTI
LAC Z PROMOTER
RIBISI
MKYLLPTAAAGLLLLAAQPAM
AQVKL
HEAVY CHAIN FRAGMENT
YPYDVPDYAS STOP
DECAPEPTIDE TAG
LEADER SEQUENCE
NOTI
XBAI
SACI
RIBISI
MKYLLPTAAAGLLLLAAQPAM
AE
LIGHT CHAIN
STOP
LEADER SEQUENCE

FIG. 7

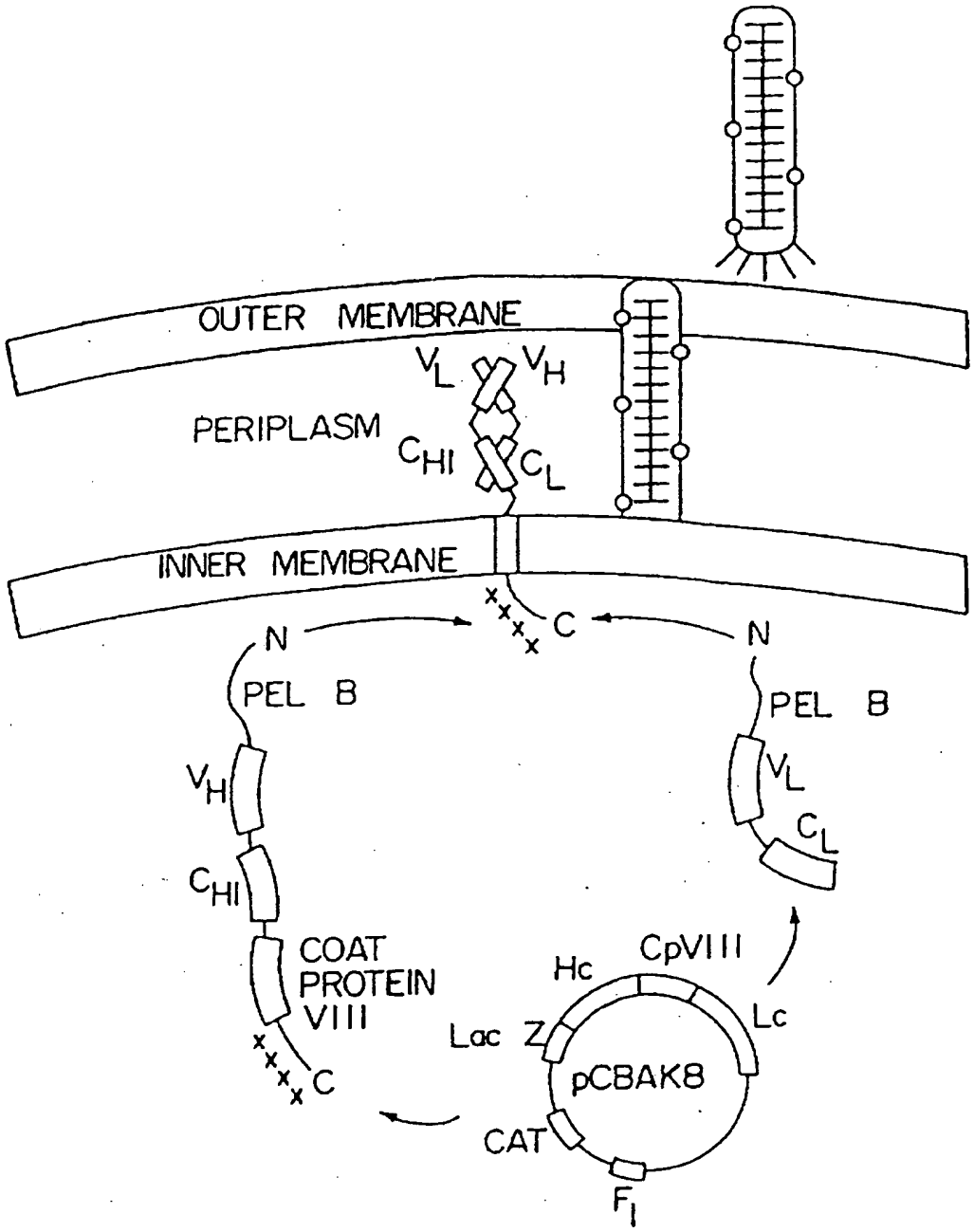

OUTER MEMBRANE
V_L
V_H
PERIPLASM
C_{HI}
C_L
INNER MEMBRANE
N
C
N
PEL B
PEL B
V_H
C_{HI}
COAT PROTEIN VIII
C
V_L
C_L
Hc
CpVIII
Lc
Lac Z
pCBAK8
CAT
F_1

FIG. 8

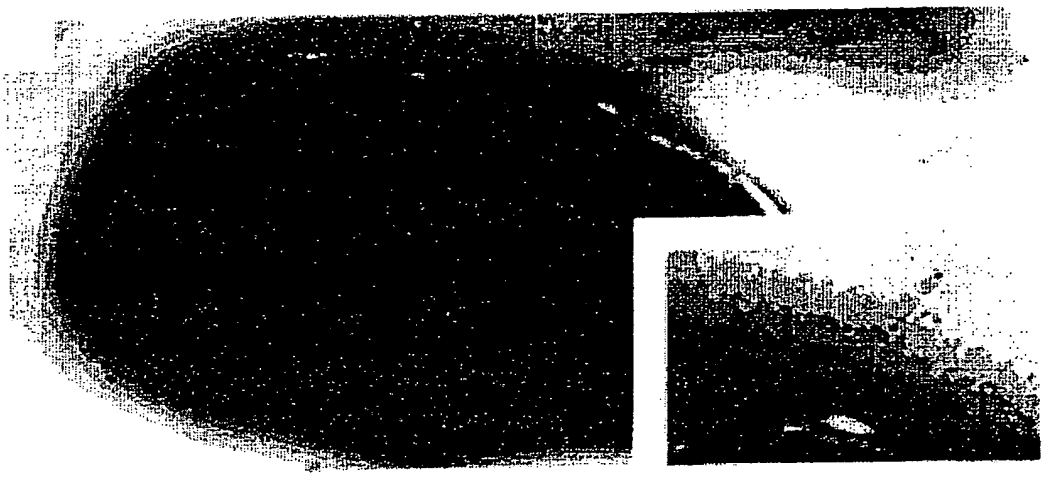


FIG. 9A-1

FIG. 9A-2

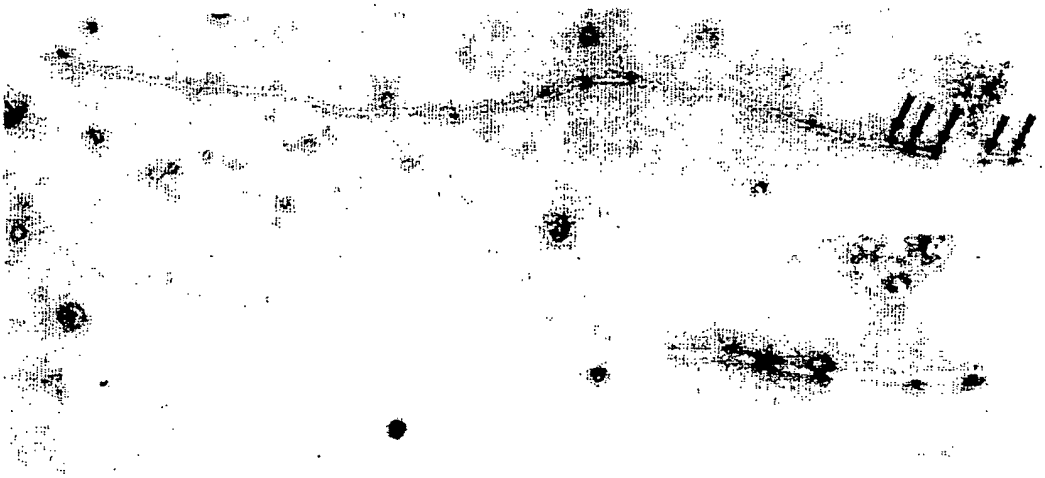


FIG. 9B-1

FIG. 9B-2

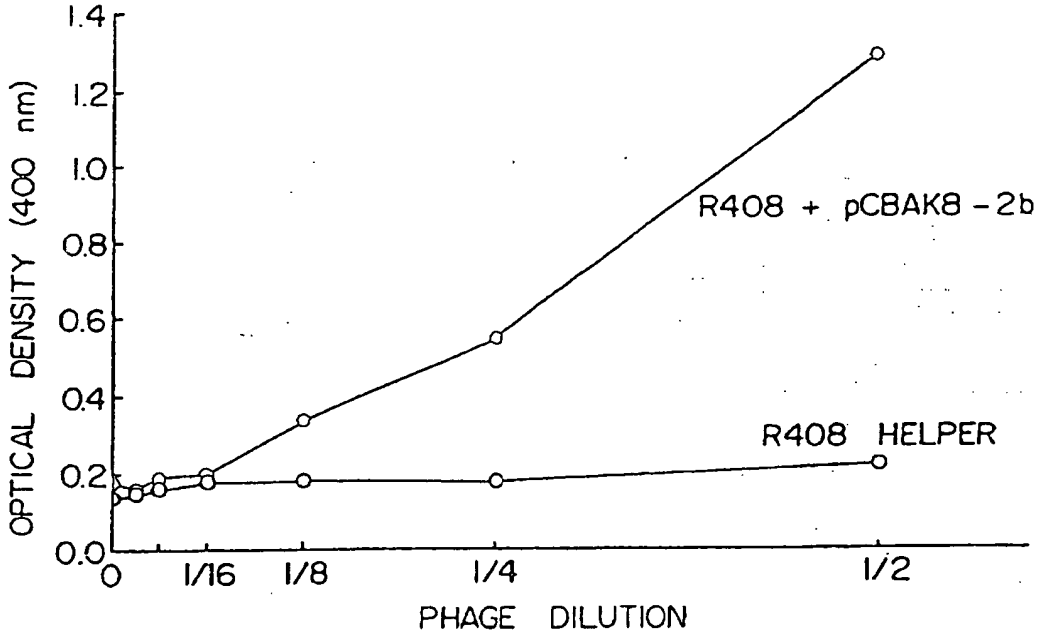

1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
OPTICAL DENSITY (400 nm)
R408 + pCBAK8 - 2b
R408 HELPER
0
1/16
1/8
1/4
1/2
PHAGE DILUTION

FIG. 10

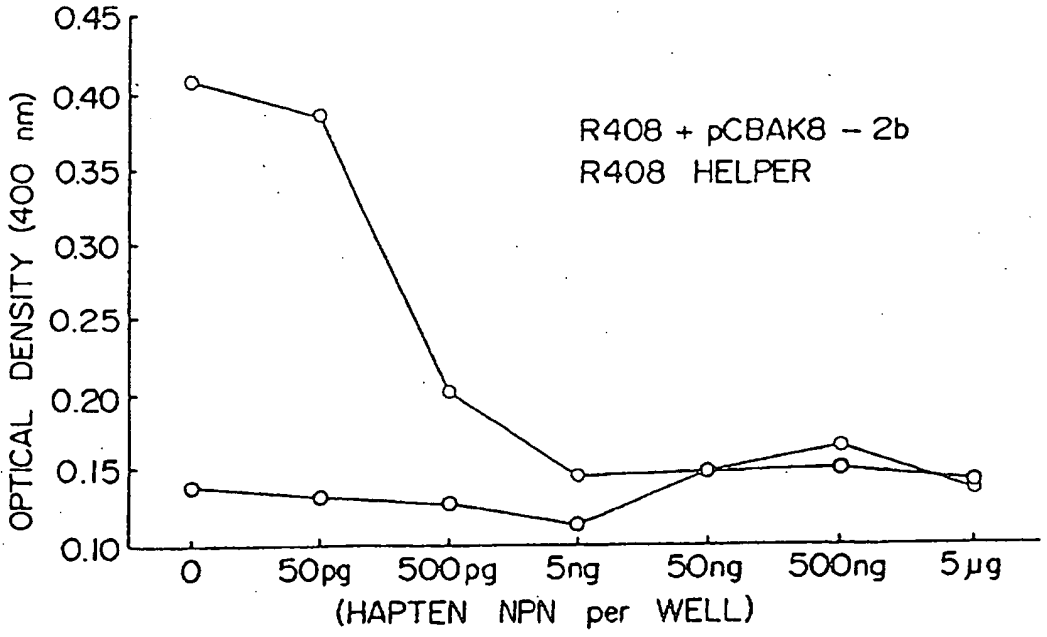

0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
OPTICAL DENSITY (400 nm)
R408 + pCBAK8 - 2b
R408 HELPER
0
50pg
500pg
5ng
50ng
500ng
5µg
(HAPTEN NPN per WELL)

FIG. 11

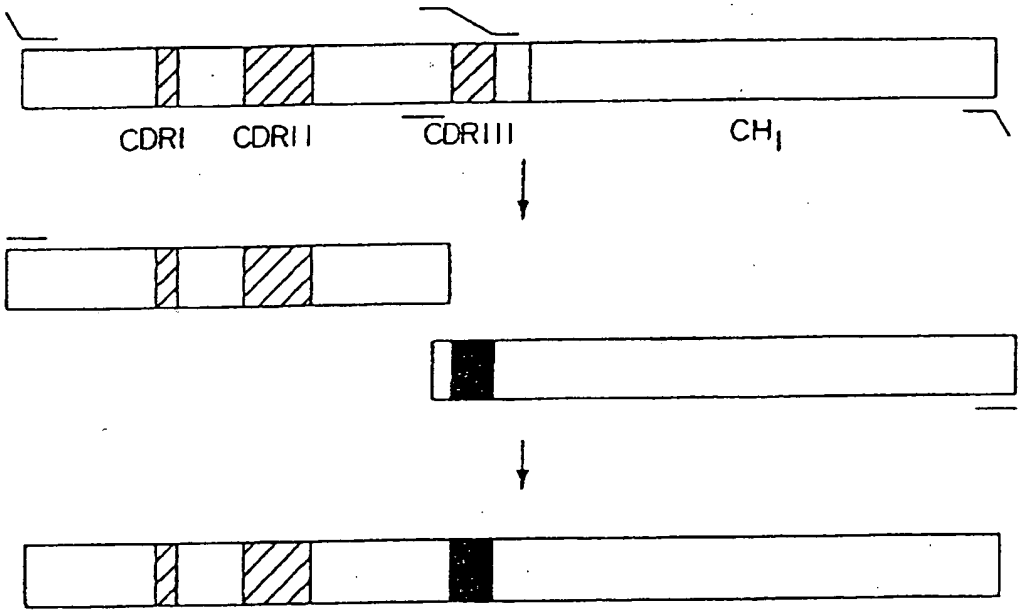


FIG. 12

HETERODIMERIC RECEPTOR LIBRARIES USING PHAGEMIDS

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation of U.S. patent application Ser. No. 07/826,623, filed on Jan. 27, 1992, now abandoned, which is a continuation-in-part of U.S. patent application Ser. No. 07/683,602, filed Apr. 10, 1991, now abandoned, the disclosures of which are hereby incorporated by reference.

This invention was made with government support under Grant No. CA 27489 awarded by the National Institutes of Health.

TECHNICAL FIELD

The present invention relates to cloning vectors and methods for producing a library of DNA molecules capable of expressing a fusion protein on the surface of a filamentous phage particle.

BACKGROUND

Filamentous bacteriophages are a group of related viruses that infect bacteria. They are termed filamentous because they are long and thin particles comprised of an elongated capsule that envelopes the deoxyribonucleic acid (DNA) that forms the bacteriophage genome. The F pili filamentous bacteriophage (Ff phage) infect only gram-negative bacteria by specifically adsorbing to the tip of F pili, and include fd, f1 and M13.

The mature capsule of Ff phage is comprised of a coat of five phage-encoded gene products: cpVIII, the major coat protein product of gene VIII that forms the bulk of the capsule; and four minor coat proteins, cpIII and cpIV at one end of the capsule and cpVII and cpIX at the other end of the capsule. The length of the capsule is formed by 2500 to 3000 copies of cpVIII in an ordered helix array that forms the characteristic filament structure. About five copies each of the minor coat proteins are present at the ends of the capsule. The gene III-encoded protein (cpIII) is typically present in 4 to 6 copies at one end of the capsule and serves as the receptor for binding of the phage to its bacterial host in the initial phase of infection. For detailed reviews of Ff phage structure, see Rasched et al., *Microbiol. Rev.*, 50:401–427 (1986); and Model et al., in "The Bacteriophages, Volume 2", R. Calendar, Ed., Plenum Press, pp. 375–456 (1988).

The assembly of a Ff phage particle involves highly complex mechanics. No phage particles are assembled within a host cell; rather, they are assembled during extrusion of the viral genome through the host cell's membrane. Prior to extrusion, the major coat protein cpVIII and the minor coat protein cpIII are synthesized and transported to the host cell's membrane. Both cpVIII and cpIII are anchored in the host cell membrane prior to their incorporation into the mature particle. In addition, the viral genome is produced and coated with cpV protein. During the extrusion process, cpV-coated genomic DNA is stripped of the cpV coat and simultaneously recoated with the mature coat proteins. The assembly mechanisms that control transferral of these proteins from the membrane to the particle is not presently known.

Both cpIII and cpVIII proteins include two domains that provide signals for assembly of the mature phage particle. The first domain is a secretion signal that directs the newly synthesized protein to the host cell membrane. The secretion signal is located at the amino terminus of the protein and targets the protein at least to the cell membrane. The second domain is a membrane anchor domain that provides signals for association with the host cell membrane and for association with the phage particle during assembly. This second signal for both cpVIII and cpIII comprises at least a hydrophobic region for spanning the membrane.

cpVIII has been extensively studied as a model membrane protein because it can integrate into lipid bilayers such as the cell membrane in an asymmetric orientation with the acidic amino terminus toward the outside and the basic carboxy terminus toward the inside of the membrane. The mature protein is about 50 amino acid residues in length of which 11 residues provide the carboxy terminus, 19 residues provide the hydrophobic transmembrane region, and the remaining residues comprise the amino terminus. Considerable research has been done on the secretion signal region of cpVIII to advance the study of membrane protein synthesis and targeting to membranes. However, little is known about the changes that are tolerated in the structure of the cpVIII membrane anchor region that would allow for assembly of phage particles.

Manipulation of the sequence of cpIII shows that the C-terminal 23 amino acid residue stretch of hydrophobic amino acids normally responsible for a membrane anchor function can be altered in a variety of ways and retain the capacity to associate with membranes. However, those anchor-modified cpIII proteins lost their ability to genetically complement gene III mutants indicating that the requirements of a membrane anchor for functional assembly have not been elucidated.

Ff phage-based expression vectors have been described in which the entire cpIII amino acid residue sequence was modified by insertion of short polypeptide "epitopes" [Parmely et al., *Gene*, 73:305–318 (1988); and Cwirla et al., *Proc. Natl. Acad. Sci. USA*, 87:6378–6382 (1990)] or an amino acid residue sequence defining a single chain antibody domain. McCafferty et al., *Science*, 348:552–554 (1990). These hybrid proteins were synthesized and assembled onto phage particles in amounts of about 5 copies per particle, a density at which normal cpIII is usually found. However, these expressed fusion proteins include the entire cpIII amino acid residue sequence and do not suggest fusion proteins that utilize only the carboxy terminal membrane anchor domain of cpIII.

In addition, no expression system has been described in which a phage coat protein has been engineered to allow assembly of a heteromeric molecule that is functional and capable of incorporation into the coat of a phage particle.

BRIEF SUMMARY OF THE INVENTION

A new, high density, surface-integration technology has been discovered for expressing a recombinant gene product on the surface of a filamentous phage containing the recombinant gene. The invention uses a filamentous phage cpVIII membrane anchor domain as a means for linking gene-product and gene during the assembly stage of filamentous phage replication.

That is, during filamentous phage replication, gene VIII-encoded proteins assemble into a matrix which encapsulates the phage genome. It has now been discovered that (1) phage assembly is not disrupted when recombinant gene VIII-encoded proteins are present, (2) recombinant gene VIII-encoded proteins can be integrated into the assembling matrix, and (3) integration into the matrix can be directed to occur in a surface-accessible orientation.

The present invention can be advantageously applied to the production of heteromeric receptors of predetermined specificity, i.e., it can be used to produce antibodies, T-cell receptors and the like that bind a preselected ligand.

Thus, the present invention provides for linking the functions of heteromeric receptor recognition and filamentous phage replication in a method for isolating a heteromeric receptor. The method produces a filamentous phage comprised of a matrix of gene VIII-encoded proteins that encapsulate a recombinant genome. The recombinant genome contains genes encoding the proteins of the heteromeric receptor proteins. The heteromeric receptor is surface-integrated into the encapsulating matrix via a gene VIII-encoded membrane anchor domain that is fused by a peptide bond during translation to one of the heteromeric receptor proteins. The heteromeric receptor and the genes which encode it are physically linked during the assembly stage of the phage replication cycle. Specifically binding the receptor-coated phage to a solid-support advantageously provides a means for isolating a recombinant genome that encodes a desired heteromeric receptor from a diverse library of recombinant genomes.

In one embodiment, the present invention contemplates an antibody molecule comprising heavy- and light-chain proteins, said heavy-chain protein comprising a V_H-domain flanked by an amino-terminal prokaryotic secretion signal domain and a carboxy-terminal filamentous phage cpVIII membrane anchor domain, said light chain protein comprising a V_L-domain fused to an amino-terminal prokaryotic secretion signal domain.

In another embodiment, the present invention contemplates a vector for expressing a fusion protein, said vector comprising a cassette that includes upstream and downstream translatable DNA sequences operatively linked via a sequence of nucleotides adapted for directional ligation of an insert DNA, said upstream sequence encoding a prokaryotic secretion signal, said downstream sequence encoding a filamentous phage gene cpVIII membrane anchor, said translatable DNA sequences operatively linked to a set of DNA expression signals for expression of said translatable DNA sequences as portions of said fusion protein.

BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings forming a portion of this disclosure:

FIG. 1 illustrates a schematic diagram of the immunoglobulin molecule showing the principal structural features. The circled area on the heavy chain represents the variable region (V_H), a polypeptide containing a biologically active (ligand binding) portion of that region, and a gene coding for that polypeptide, are produced by the methods of the present invention.

FIG. 2A is a diagrammatic sketch of a heavy (H) chain of human IgG (IgG1 subclass). Numbering is from the N-terminus on the left to the C-terminus on the right. Note the presence of four domains, each containing an intrachain disulfide bond (S—S) spanning approximately 60 amino acid residues. The symbol CHO stands for carbohydrate. The V region of the heavy (H) chain (V_H) resembles V_L in having three hypervariable CDR (not shown).

FIG. 2B-1 is a diagrammatic sketch of a human light (Kappa) chain (Panel 1). Numbering is from the N-terminus on the left to the C-terminus on the right. Note the intrachain disulfide bond (S—S) spanning about the same number of amino acid residues in the V_L and C_L domains.

FIG. 2B-2 shows the locations of the complementarity-determining regions (CDR) in the V_L domain. Segments outside the CDR are the framework segments (FR).

FIG. 3 illustrates the sequence of the double-stranded synthetic DNA inserted into Lambda Zap to produce a Lambda Hc2 expression vector. The preparation of the double-stranded synthetic DNA insert is described in Example 1a(ii). The various features required for this vector to express the V_H-coding DNA homologs include the Shine-Dalgarno ribosome binding site, a leader sequence to direct the expressed protein to the periplasm as described by Mouva et al., *J. Biol. Chem.*, 255:27, 1980, and various restriction enzyme sites used to operatively link the V_H homologs to the expression vector. The V_H expression vector sequence also contains a short nucleic acid sequence that codes for amino acids typically found in variable regions heavy chain (V_H Backbone). This V_H Backbone is just upstream and in the proper reading as the V_H DNA homologs that are operatively linked into the Xho I and Spe I cloning sites. The sequences of the top and bottom strands of the double-stranded synthetic DNA insert are listed respectively as SEQ. ID. NO. 1 and SEQ. ID. NO. 2. The synthetic DNA insert is directionally ligated into Lambda Zap II digested with the restriction enzymes Not 1 and Xho I to form Lambda Hc2 expression vector.

FIG. 4 illustrates the major features of the bacterial expression vector Lambda Hc2 (V_H expression vector). The synthetic DNA sequence from FIG. 3 is shown at the top along with the T_3 polymerase promoter from Lambda Zap II. The orientation of the insert in Lambda Zap II is shown. The V_H DNA homologs are inserted into the Xho I and Spe I cloning sites. The read through transcription produces the decapeptide epitope (tag) that is located just 3' of the cloning site.

FIG. 5 illustrates the sequence of the double-stranded synthetic DNA inserted into Lambda Zap to produce a Lambda Lc2 expression vector. The various features required for this vector to express the V_L-coding DNA homologs are described in FIG. 3. The V_L-coding DNA homologs are operatively linked into the Lc2 sequence at the Sac I and Xho I restriction sites. The sequences of the top and bottom strands of the double-stranded synthetic DNA insert are listed respectively as SEQ. ID. NO. 3 and SEQ. ID. NO. 4. The synthetic DNA insert is directionally ligated into Lambda Zap II digested with the restriction enzymes Sac I and Not I to form Lambda Lc2 expression vector.

FIG. 6 illustrates the major features of the bacterial expression vector Lc2 (V_L expression vector). The synthetic DNA sequence from FIG. 5 is shown at the top along with the T_3 polymerase promoter from Lambda Zap II. The orientation of the insert in Lambda Zap II is shown. The V_L DNA homologs are inserted into the Sac I and Xho I cloning sites.

FIG. 7 illustrates the dicistronic expression vector, pComb, in the form of a phagemid expression vector. To produce pComb, phagemids were first excised from the expression vectors, Lambda Hc2 and Lambda Lc2, using an in vivo excision protocol according to manufacturers instructions (Stratagene, La Jolla, Calif.). The pComb expression vector is prepared from Lambda Hc2 and Lambda Lc2 which do not contain V_H-coding or V_L-coding DNA homologs. The in vivo excision protocol moved the cloned insert from the Lambda Hc2 and Lc2 vectors into a phagemid vector. The resultant phagemids contained the same nucleotide sequences for antibody fragment cloning and expression as did the parent vectors. Hc2 and Lc2 phagemid expression vectors were separately restriction digested with Sca I and EcoR I. The linearized phagemids were ligated via the Sca I and EcoR I cohesive termini to form the dicistronic (combinatorial) vector, pComb.

FIG. 8 illustrates a schematic diagram of the composition of pCBAK8-2b phagemid vector, the pathway for Fab assembly and incorporation in phage coat. The vector carries the chloramphenicol acetyl transferase (CAT) marker gene in addition to the nucleotide residue sequences encoding the Fd-cpVIII fusion protein and the kappa chain. The f1 phage origin of replication facilitates the generation of single stranded phagemid. The isopropyl thiogalactopyranoside (IPTG) induced expression of a dicistronic message encoding the Fd-cpVIII fusion (V_H, C_{H1}, cpVIII) and the light chain (V_L, C_L) leads to the formation of heavy and light chains. Each chain is delivered to the periplasmic space by the pelB target sequence, which is subsequently cleaved. The heavy chain is anchored in the membrane by cpVIII fusion while the light chain is secreted into the periplasm. The heavy chain in the presence of light chain assembles to form Fab molecules. The Fabs are incorporated into phage particles via cpVIII (black dots).

FIG. 9 illustrates the electron micrographic localization of 5–7 nm colloidal gold particles coated with NPN-BSA conjugate along the surface of filamentous phage, and from phage emerging from a bacterial cell. Panel 9A shows filamentous phage emerging from the surface of the bacterial cell specifically labelled with the colloidal gold particles coated with BSA-NPN antigen. Panel 9B shows a portion of a mature filamentous phage on the length of which is exhibited the labelling of antigen binding sites.

FIG. 10 illustrates the results of a two-site ELISA for assaying for the presence and function of Fab antibody attached to the surface of bacteriophage particles as described in Example 4b. For expression of Fab antibody on phage surfaces, XL1-Blue cells were transformed with the phagemid expression vector, pCBAK8-2b. The inducer, isopropyl thiogalactopyranoside (IPTG), was admixed with the bacterial suspension at a final concentration of 1 mM for one hour. Helper phage was then admixed with the bacterial suspension to initiate the generation of copies of the sense strand of the phagemid DNA. After a two hour maintenance period, bacterial supernatants containing bacteriophage particles were collected for assaying in ELISA.

Specific titratable binding of NPN-Fab-expressing bacteriophage particles to NPN-coated plates was exhibited. No binding was detected with helper phage alone.

FIG. 11 illustrates the inhibition of NPN-Fab expressing bacteriophage to NPN antigen-coated plates with the addition of increasing amounts of free hapten. The assays were performed as described in FIG. 10. Complete inhibition of binding was observed with 5 ng of added free NPN hapten.

FIG. 12 illustrates schematically the process of mutagenizing the CDR3 region of a heavy chain fragment resulting in an alteration of binding specificity. The oligonucleotide primers are indicated by black bars. The process is described in Example 6.

DETAILED DESCRIPTION OF THE INVENTION

A. Definitions

Amino Acid Residue: An amino acid formed upon chemical digestion (hydrolysis) of a polypeptide at its peptide linkages. The amino acid residues described herein are preferably in the "L" isomeric form. However, residues in the "D" isomeric form can be substituted for any L-amino acid residue, as long as the desired functional property is retained by the polypeptide. NH_2 refers to the free amino group present at the amino terminus of a polypeptide. COOH refers to the free carboxy group present at the carboxy terminus of a polypeptide. In keeping with standard polypeptide nomenclature (described in *J. Biol. Chem.*, 243:3552–59 (1969) and adopted at 37 C.F.R. 1.822(b)(2)), abbreviations for amino acid residues are shown in the following Table of Correspondence:

TABLE OF CORRESPONDENCE

SYMBOL		
1-Letter	3-Letter	AMINO ACID
Y	Tyr	tyrosine
G	Gly	glycine
F	Phe	phenylalanine
M	Met	methionine
A	Ala	alanine
S	Ser	serine
I	Ile	isoleucine
L	Leu	leucine
T	Thr	threonine
V	Val	valine
P	Pro	proline
K	Lys	lysine
H	His	histidine
Q	Gln	glutamine
E	Glu	glutamic acid
Z	Glx	Glu and/or Gln
W	Trp	tryptophan
R	Arg	arginine
D	Asp	aspartic acid
N	Asn	asparagine
B	Asx	Asn and/or Asp
C	Cys	cysteine
J	Xaa	Unknown or other

It should be noted that all amino acid residue sequences represented herein by formulae have a left-to-right orientation in the conventional direction of amino terminus to carboxy terminus. In addition, the phrase "amino acid residue" is broadly defined to include the amino acids listed in the Table of Correspondence and modified and unusual amino acids, such as those listed in 37 C.F.R. 1.822(b)(4), and incorporated herein by reference. Furthermore, it should be noted that a dash at the beginning or end of an amino acid residue sequence indicates a peptide bond to a further sequence of one or more amino acid residues or a covalent bond to an amino-terminal group such as NH_2 or acetyl or to a carboxy-terminal group such as COOH.

Nucleotide: A monomeric unit of DNA or RNA consisting of a sugar moiety (pentose), a phosphate, and a nitrogenous heterocyclic base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of the pentose) and that combination of base and sugar is a nucleoside. When the nucleoside contains a phosphate group bonded to the 3' or 5' position of the pentose it is referred to as a nucleotide. A sequence of operatively linked nucleotides is typically referred to herein as a "base sequence" or "nucleotide sequence", and their grammatical equivalents, and is represented herein by a formula whose left to right orientation is in the conventional direction of 5'-terminus to 3'-terminus.

Base Pair (bp): A partnership of adenine (A) with thymine (T), or of cytosine (C) with guanine (G) in a double stranded DNA molecule. In RNA, uracil (U) is substituted for thymine.

Nucleic Acid: A polymer of nucleotides, either single or double stranded.

Polynucleotide: a polymer of single or double stranded nucleotides. As used herein "polynucleotide" and its grammatical equivalents will include the full range of nucleic acids. A polynucleotide will typically refer to a nucleic acid

molecule comprised of a linear strand of two or more deoxyribonucleotides and/or ribonucleotides. The exact size will depend on many factors, which in turn depends on the ultimate conditions of use, as is well known in the art. The polynucleotides of the present invention include primers, probes, RNA/DNA segments, oligonucleotides or "oligos" (relatively short polynucleotides), genes, vectors, plasmids, and the like.

Gene: A nucleic acid whose nucleotide sequence codes for an RNA or polypeptide. A gene can be either RNA or DNA.

Duplex DNA: a double-stranded nucleic acid molecule comprising two strands of substantially complementary polynucleotides held together by one or more hydrogen bonds between each of the complementary bases present in a base pair of the duplex. Because the nucleotides that form a base pair can be either a ribonucleotide base or a deoxyribonucleotide base, the phrase "duplex DNA" refers to either a DNA-DNA duplex comprising two DNA strands (ds DNA), or an RNA-DNA duplex comprising one DNA and one RNA strand.

Complementary Bases: Nucleotides that normally pair up when DNA or RNA adopts a double stranded configuration.

Complementary Nucleotide Sequence: A sequence of nucleotides in a single-stranded molecule of DNA or RNA that is sufficiently complementary to that on another single strand to specifically hybridize to it with consequent hydrogen bonding.

Conserved: A nucleotide sequence is conserved with respect to a preselected (reference) sequence if it non-randomly hybridizes to an exact complement of the preselected sequence.

Hybridization: The pairing of substantially complementary nucleotide sequences (strands of nucleic acid) to form a duplex or heteroduplex by the establishment of hydrogen bonds between complementary base pairs. It is a specific, i.e. non-random, interaction between two complementary polynucleotides that can be competitively inhibited.

Nucleotide Analog: A purine or pyrimidine nucleotide that differs structurally from A, T, G, C, or U, but is sufficiently similar to substitute for the normal nucleotide in a nucleic acid molecule.

DNA Homolog: Is a nucleic acid having a preselected conserved nucleotide sequence and a sequence coding for a receptor capable of binding a preselected ligand.

Recombinant DNA (rDNA) molecule: a DNA molecule produced by operatively linking two DNA segments. Thus, a recombinant DNA molecule is a hybrid DNA molecule comprising at least two nucleotide sequences not normally found together in nature. rDNA's not having a common biological origin, i.e., evolutionarily different, are said to be "heterologous".

Vector: a rDNA molecule capable of autonomous replication in a cell and to which a DNA segment, e.g., gene or polynucleotide, can be operatively linked so as to bring about replication of the attached segment. Vectors capable of directing the expression of genes encoding for one or more proteins are referred to herein as "expression vectors". Particularly important vectors allow cloning of cDNA (complementary DNA) from mRNAs produced using reverse transcriptase.

Receptor: A receptor is a molecule, such as a protein, glycoprotein and the like, that can specifically (non-randomly) bind to another molecule.

Antibody: The term antibody in its various grammatical forms is used herein to refer to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antibody combining site or paratope. Exemplary antibody molecules are intact immunoglobulin molecules, substantially intact immunoglobulin molecules and portions of an immunoglobulin molecule, including those portions known in the art as Fab, Fab', $F(ab')_2$ and F(v).

Antibody Combining Site: An antibody combining site is that structural portion of an antibody molecule comprised of a heavy and light chain variable and hypervariable regions that specifically binds (immunoreacts with) an antigen. The term immunoreact in its various forms means specific binding between an antigenic determinant-containing molecule and a molecule containing an antibody combining site such as a whole antibody molecule or a portion thereof.

Monoclonal Antibody: The phrase monoclonal antibody in its various grammatical forms refers to a population of antibody molecules that contains only one species of antibody combining site capable of immunoreacting with a particular antigen. A monoclonal antibody thus typically displays a single binding affinity for any antigen with which it immunoreacts. A monoclonal antibody may therefore contain an antibody molecule having a plurality of antibody combining sites, each immunospecific for a different antigen, e.g., a bispecific monoclonal antibody.

Fusion Protein: A protein comprised of at least two polypeptides and a linking sequence to operatively link the two polypeptides into one continuous polypeptide. The two polypeptides linked in a fusion protein are typically derived from two independent sources, and therefore a fusion protein comprises two linked polypeptides not normally found linked in nature.

Upstream: In the direction opposite to the direction of DNA transcription, and therefore going from 5' to 3' on the non-coding strand, or 3' to 5' on the mRNA.

Downstream: Further along a DNA sequence in the direction of sequence transcription or read out, that is traveling in a 3'- to 5'-direction along the non-coding strand of the DNA or 5'- to 3'-direction along the RNA transcript.

Cistron: Sequence of nucleotides in a DNA molecule coding for an amino acid residue sequence and including upstream and downstream DNA expression control elements.

Stop Codon: Any of three codons that do not code for an amino acid, but instead cause termination of protein synthesis. They are UAG, UAA and UGA and are also referred to as a nonsense or termination codon.

Leader Polypeptide: A short length of amino acid sequence at the amino end of a protein, which carries or directs the protein through the inner membrane and so ensures its eventual secretion into the periplasmic space and perhaps beyond. The leader sequence peptide is commonly removed before the protein becomes active.

Reading Frame: Particular sequence of contiguous nucleotide triplets (codons) employed in translation. The reading frame depends on the location of the translation initiation codon.

B. Filamentous Phage

The present invention contemplates a filamentous phage comprising a matrix of cpVIII proteins encapsulating a genome encoding first and second polypeptides. The phage further comprises a heteromeric receptor comprised of the first and second polypeptides surface-integrated into the matrix via a cpVIII membrane anchor domain fused to at

least one of the first or second polypeptides. Preferably, the first and second polypeptides are V_H and V_L proteins, respectively.

The first and second polypeptides are capable of autogenous assembly into a functional receptor, which is expressed on the outer surface in a manner accessible to ligand, i.e. they are surface-integrated into the phage. Typically, the receptor is comprised of a linking polypeptide that contains the cpVIII membrane anchor domain, such as a polypeptide described in Section C, and a non-linking polypeptide(s).

Because the receptor is linked to the phage in a surface accessible manner, the phage can be advantageously used as a solid-phase affinity sorbent. In preferred embodiments, the phage are linked, preferably removably linked, to a solid (aqueous insoluble) matrix such as agarose, cellulose, synthetic reins, polysaccharides and the like. For example, transformants shedding the phage can be applied to and retained in a column and maintained under conditions that support shedding of the phage. An aqueous composition containing a ligand that binds to the receptor expressed by the phage is then passed through the column at a predetermined rate and under receptor-binding conditions to form a solid-phase receptor-ligand complex. The column is then washed to remove unbound material, leaving the ligand bound to the solid-phase phage. The ligand can then be removed and recovered by washing the column with a buffer that promotes dissociation of the receptor-ligand complex.

Alternatively, purified phage can be admixed with a aqueous solution containing the ligand to be affinity purified. The receptor/ligand binding reaction admixture thus formed is maintained for a time period and under binding conditions sufficient for a phage-linked receptor-ligand complex to form. The phage-bound ligand (ligand-bearing phage) are then separated and recovered from the unbound materials, such as by centrifugation, electrophoresis, precipitation, and the like.

C. DNA Expression Vectors

A vector of the present invention is a recombinant DNA (rDNA) molecule adapted for receiving and expressing translatable DNA sequences in the form of a fusion protein containing a filamentous phage gene VIII membrane anchor domain and a prokaryotic secretion signal domain. The vector comprises a cassette that includes upstream and downstream translatable DNA sequences operatively linked via a sequence of nucleotides adapted for directional ligation. The upstream translatable sequence encodes the secretion signal. The downstream translatable sequence encodes the filamentous phage membrane anchor. The cassette preferably includes DNA expression control sequences for expressing the fusion protein that is produced when a translatable DNA sequence is directionally inserted into the cassette via the sequence of nucleotides adapted for directional ligation.

An expression vector is characterized as being capable of expressing, in a compatible host, a structural gene product such as a fusion protein of the present invention.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting between different genetic environments another nucleic acid to which it has been operatively linked. Preferred vectors are those capable of autonomous replication and expression of structural gene products present in the DNA segments to which they are operatively linked.

As used herein with regard to DNA sequences or segments, the phrase "operatively linked" means the sequences or segments have been covalently joined into one shard of DNA, whether in single or double stranded form.

The choice of vector to which a cassette of this invention is operatively linked depends directly, as is well known in the art, on the functional properties desired, e.g., vector replication and protein expression, and the host cell to be transformed, these being limitations inherent in the art of constructing recombinant DNA molecules.

In preferred embodiments, the vector utilized includes a prokaryotic replicon i.e., a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extra chromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, those embodiments that include a prokaryotic replicon also include a gene whose expression confers a selective advantage, such as drug resistance, to a bacterial host transformed therewith. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline. Vectors typically also contain convenient restriction sites for insertion of translatable DNA sequences. Exemplary vectors are the plasmids pUC8, pUC9, pBR322, and pBR329 available from BioRad Laboratories, (Richmond, Calif.) and pPL and pKK223 available from Pharmacia, (Piscataway, N.J.).

A sequence of nucleotides adapted for directional ligation, i.e., a polylinker, is a region of the DNA expression vector that (1) operatively links for replication and transport the upstream and downstream translatable DNA sequences and (2) provides a site or means for directional ligation of a DNA sequence into the vector. Typically, a directional polylinker is a sequence of nucleotides that defines two or more restriction endonuclease recognition sequences, or restriction sites. Upon restriction cleavage, the two sites yield cohesive termini to which a translatable DNA sequence can be ligated to the DNA expression vector. Preferably, the two restriction sites provide, upon restriction cleavage, cohesive termini that are non-complementary and thereby permit directional insertion of a translatable DNA sequence into the cassette. In one embodiment, the directional ligation means is provided by nucleotides present in the upstream translatable DNA sequence, downstream translatable DNA sequence, or both. In another embodiment, the sequence of nucleotides adapted for directional ligation comprises a sequence of nucleotides that defines multiple directional cloning means. Where the sequence of nucleotides adapted for directional ligation defines numerous restriction sites, it is referred to as a multiple cloning site.

A translatable DNA sequence is a linear series of nucleotides that provide an uninterrupted series of at least 8 codons that encode a polypeptide in one reading frame.

An upstream translatable DNA sequence encodes a prokaryotic secretion signal. The secretion signal is a leader peptide domain of protein that targets the protein to the periplasmic membrane of gram negative bacteria.

A preferred secretion signal is a pelB secretion signal. The predicted amino acid residue sequences of the secretion signal domain from two pelB gene product variants from *Erwinia carotova* are shown in Table 1 as described by Lei, et al., *Nature*, 331:543–546 (1988). A particularly preferred pelB secretion signal is also shown in Table 1.

The leader sequence of the pelB protein has previously been used as a secretion signal for fusion proteins. Better et al., *Science*, 240:1041–1043 (1988); Sastry et al., *Proc. Natl. Acad. Sci. USA*, 86:5728–5732 (1989); and Mullinax et al., *Proc. Natl. Acad. Sci. USA*, 87:8095–8099 (1990).

Amino acid residue sequences for other secretion signal polypeptide domains from *E. coli* useful in this invention are also listed in Table 1. Oliver, In Neidhard, F. C. (ed.). *Escherichia coli* and *Salmonella Typhimurium*, American Society for Microbiology, Washington, D.C., 1:56–69 (1987).

A translatable DNA sequence encoding the pelB secretion signal having the amino acid residue sequence shown in SEQ. ID. NO. 5 is a preferred DNA sequence for inclusion in a DNA expression vector of this invention.

TABLE 1

Leader Sequences

SEQ. ID. NO.	Type	Amino Acid Residue Sequence
(5)	PelB[1]	MetLysTyrLeuLeuProThrAlaAlaAlaGlyLeuLeu LeuLeuAlaAlaGlnProAlaMet
(6)	PelB[2]	MetLysTyrLeuLeuProThrAlaAlaAlaGlyLeuLeu LeuLeuAlaAlaGlnProAlaGlnProAlaMetAla
(7)	PelB[3]	MetLysSerLeuIleThrProIleAlaAlaGlyLeuLeu LeuAlaPheSerGlnTyrSerLeuAla
(8)	MalE[4]	MetLysIleLysThrGlyAlaArgIleLeuAlaLeuSer AlaLeuThrThrMetMetPheSerAlaSerAlaLeuAla LysIle
(9)	OmpF[4]	MetMetLysArgAsnIleLeuAlaValIleValProAla LeuLeuValAlaGlyThrAlaAsnAlaAlaGlu
(10)	PhoA[4]	MetLysGlnSerThrIleAlaLeuAlaLeuLeuProLeu LeuPheThrProValThrLysAlaArgThr
(11)	Bla[4]	MetSerIleGlnHisPheArgValAlaLeuIleProPhe PheAlaAlaPheCysLeuProValPheAlaHisPro
(12)	LamB[4]	MetMetIleThrLeuArgLysLeuProLeuAlaValAla ValAlaAlaGlyValMetSerAlaGlnAlaMetAlaVal Asp
(13)	Lpp[4]	MetLysAlaThrLysLeuValLeuGlyAlaValIleLeu GlySerThrLeuLeuAlaGlyCysSer
(14)	cpVIII[5]	MetLysLysSerLeuValLeuLysAlaSerValAlaVal AlaThrLeuValProMetLeuSerPheAla
(15)	cpIII[6]	MetLysLysLeuLeuPheAlaIleProLeuValValPro PheTyrSerHisSer

[1]pelB used in this invention
[2]pelB from *Erwinia carotovora* gene
[3]pelB from *Erwinia carotovora* EC 16 gene
[4]leader sequences from *E. coli*
[5]leader sequence for cpVIII
[6]leader sequence for cpIII

A downstream translatable DNA sequence encodes a filamentous phage membrane anchor. Preferred membrane anchors are obtainable from filamentous phage M13, f1, fd, and the like. Preferred membrane anchor domains are found in the coat proteins encoded by gene III and gene VIII. Thus, a downstream translatable DNA sequence encodes an amino acid residue sequence that corresponds, and preferably is identical, to the membrane anchor domain of either a filamentous phage gene III or gene VIII coat protein.

The membrane anchor domain of a filamentous phage coat protein is the carboxy terminal region of the coat protein and includes a region of hydrophobic amino acid residues for spanning a lipid bilayer membrane, and a region of charged amino acid residues normally found at the cytoplasmic face of the membrane and extending away from the membrane. In the phage f1, gene VIII coat protein's membrane spanning region comprises residue Trp-26 through Lys-40, and the cytoplasmic region comprises the carboxy-terminal 11 residues from 41 to 52. Ohkawa et al., *J. Biol. Chem.*, 256:9951–9958 (1981).

The amino acid residue sequence of a preferred membrane anchor domain derived from the M13 filamentous phage gene III coat protein (also designated cpIII) has a sequence shown in SEQ. ID. NO. 16 from residue 1 to residue 211. Gene III coat protein is present on a mature filamentous phage at one end of the phage particle with typically about 4 to 6 copies of the coat protein.

The amino acid residue sequence of a preferred membrane anchor domain derived from the M13 filamentous phage gene VIII coat protein (also designated cpVIII) has a sequence shown in SEQ. ID. NO. 17 from residue 1 to residue 50. Gene VIII coat protein is present on a mature filamentous phage over the majority of the phage particle with typically about 2500 to 3000 copies of the coat protein.

For detailed descriptions of the structure of filamentous phage particles, their coat proteins and particle assembly, see the reviews by Rached et al., *Microbiol. Rev.*, 50:401–427 (1986); and Model et al., in "The Bacteriophages: Vol. 2", R. Calendar, ed. Plenum Publishing Co., pp. 375–456, (1988).

A cassette in a DNA expression vector of this invention is the region of the vector that forms, upon insertion of a translatable DNA sequence, a sequence of nucleotides capable of expressing, in an appropriate host, a fusion protein of this invention. The expression-competent sequence of nucleotides is referred to as a cistron. Thus, the cassette comprises DNA expression control elements operatively linked to the upstream and downstream translatable DNA sequences. A cistron is formed when a translatable DNA sequence is directionally inserted (directionally ligated) between the upstream and downstream sequences via the sequence of nucleotides adapted for that purpose. The resulting three translatable DNA sequences, namely the upstream, the inserted and the downstream sequences, are all operatively linked in the same reading frame.

DNA expression control sequences comprise a set of DNA expression signals for expressing a structural gene product and include both 5' and 3' elements, as is well known, operatively linked to the cistron such that the cistron is able to express a structural gene product. The 5' control sequences define a promoter for initiating transcription and a ribosome binding site operatively linked at the 5' terminus of the upstream translatable DNA sequence.

To achieve high levels of gene expression in *E. coli*, it is necessary to use not only strong promoters to generate large quantities of mRNA, but also ribosome binding sites to ensure that the mRNA is efficiently translated. In *E. coli*, the ribosome binding site includes an initiation codon (AUG) and a sequence 3–9 nucleotides long located 3–11 nucleotides upstream from the initiation codon [Shine et al., *Nature*, 254:34 (1975)] The sequence, AGGAGGU, which is called the Shine-Dalgarno (SD) sequence, is complementary to the 3' end of *E. coli* 16S mRNA. Binding of the ribosome to mRNA and the sequence at the 3' end of the mRNA can be affected by several factors:

(i) The degree of complementarity between the SD sequence and 3' end of the 16S tRNA.

(ii) The spacing and possibly the DNA sequence lying between the SD sequence and the AUG [Roberts et al., *Proc. Natl. Acad. Sci. USA*, 76:760 (1979a); Roberts et al., *Proc. Natl. Acad. Sci. USA*, 76:5596 (1979b); Guarente et al., *Science*, 209:1428 (1980); and Guarente et al., *Cell*, 20:543 (1980).] Optimization is achieved by measuring the level of expression of genes in plasmids in which this spacing is systematically altered. Comparison of different mRNAs shows that there are statistically preferred sequences from positions −20 to +13 (where the A of the AUG is position 0) [Gold et al., *Annu. Rev. Microbiol.*, 35:365 (1981)]. Leader sequences have been shown to influence translation dramatically (Roberts et al., 1979 a, b supra).

(iii) The nucleotide sequence following the AUG, which affects ribosome binding [Taniguchi et al., *J. Mol. Biol.*, 118:533 (1978)].

Useful ribosome binding sites are shown in Table 2 below.

TABLE 2

Ribosome Binding Sites*

SEQ. ID. NO.		
1.	(18)	5' AAUCUUGGAGGCUUUUUUAUGGUUCGUUCU
2.	(19)	5' UAACUAAGGAUGAAAUGCAUGUCUAAGACA
3.	(20)	5' UCCUAGGAGGUUUGACCUAUGCGAGCUUUU
4.	(21)	5' AUGUACUAAGGAGGUUGUAUGGAACAACGC

*Sequences of initiation regions for protein synthesis in four phage mRNA molecules are underlined.
AUG = initiation codon (double underlined)
1. = Phage ϕX174 gene-A protein
2. = Phage Qβ replicase
3. = Phage R17 gene-A protein
4. = Phage lambda gene-cro protein

The 3' control sequences define at least one termination (stop) codon in frame with and operatively linked to the downstream translatable DNA sequence.

Thus, a DNA expression vector of this invention provides a system for cloning translatable DNA sequences into the cassette portion of the vector to produce a cistron capable of expressing a fusion protein of this invention.

In preferred embodiments, a DNA expression vector provides a system for independently cloning two translatable DNA sequences into two separate cassettes present in the vector, to form two separate cistrons for expressing both polypeptides of a heterodimeric receptor, or the ligand binding portions of the polypeptides that comprise a heterodimeric receptor. The DNA expression vector for expressing two cistrons is referred to as a dicistronic expression vector.

Thus, a preferred DNA expression vector of this invention comprises, in addition to the cassette previously described in detail, a second cassette for expressing a second fusion protein. The second cassette includes a third translatable DNA sequence that encodes a secretion signal, as defined herein before, operatively linked at its 3' terminus via a sequence of nucleotides adapted for directional ligation to a downstream DNA sequence defining at least one stop codon. The third translatable DNA sequence is operatively linked at its 5' terminus to DNA expression control sequences forming the 5' elements defined above. The second cassette is capable, upon insertion of a translatable DNA sequence, of expressing the second fusion protein.

In a preferred embodiment, a DNA expression vector is designed for convenient manipulation in the form of a filamentous phage particle according to the teachings of the present invention. In this embodiment, a DNA expression vector further contains a nucleotide sequence that defines a filamentous phage origin of replication such that the vector, upon presentation of the appropriate genetic complementation, can replicate as a filamentous phage in single stranded replicative form and be packaged into filamentous phage particles. This feature provides the ability of the DNA expression vector to be packaged into phage particles for subsequent segregation of the particle, and vector contained therein, away from other particles that comprise a population of phage particles.

A filamentous phage origin of replication is a region of the phage genome, as is well known, that defines sites for initiation of replication, termination of replication and packaging of the replicative form produced by replication. See, for example, Rasched et al., *Microbiol. Rev.*, 50:401–427 (1986); and Horiuchi, *J. Mol. Biol.*, 188:215–223 (1986).

A preferred filamentous phage origin of replication for use in the present invention is a M13, f1 or fd phage origin of replication. Particularly preferred is a filamentous phage origin of replication having a sequence shown in SEQ. ID. NO. 117 and described by Short et al., *Nucl. Acids Res.*, 16:7583–7600 (1988). Preferred DNA expression vectors are the dicistronic expression vectors pCOMB8 and pCOMB3 described in Example 1b(i) and 1b(ii), respectively.

D. Polypeptides

In another embodiment, the present invention contemplates a polypeptide comprising an insert domain flanked by an amino-terminal secretion signal domain and a carboxy-terminal filamentous phage coat protein membrane anchor domain.

Preferably, the polypeptide is a fusion protein having a receptor domain comprised of an amino acid residue sequence that defines the ligand binding domain of a receptor protein positioned between a prokaryotic secretion signal domain and a gene VIII-encoded (cpVIII) membrane anchor domain. In preferred embodiments, the receptor protein is a polypeptide chain of a heterodimeric receptor. Insofar as the polypeptide has a receptor domain, it is also referred to herein as a receptor. In other preferred embodiments the secretion signal domain is a pelB secretion signal as described herein.

Preferred heterodimeric receptors include immunoglobulins, major histocompatibility antigens of class I or II, lymphocyte receptors, integrins and the like heterodimeric receptors.

In one embodiment, a polypeptide of this invention has an amino acid residue sequence that can be represented by the formula, shown in the direction of amino- to carboxy terminus:

$$NH_2—O—(U)_m—V—(X)_n—Z—COOH, \quad (F1)$$

where O represents an amino acid residue sequence defining a secretion signal, U represents a first spacer polypeptide, V represents an amino acid residue sequence defining a receptor domain, X represents a second spacer polypeptide, and Z represents an amino acid residue sequence defining a filamentous phage coat protein membrane anchor, with the proviso that m is the integer 0 or 1 such that when m is O, U is not present, and when m is 1, U is present, and n is 0 or 1 such that when n is O, X is not present and when n is 1, X is present.

In the formula (F1), the secretion signal and the filamentous phage coat protein membrane anchor are as defined herein above. Particularly preferred is a polypeptide according to formula (F1) where Z defines the gene VIII membrane anchor as described herein. In another preferred embodiment the secretion signal is the pelB secretion signal.

In one embodiment, V is an amino acid residue sequence that defines the ligand binding domain of a chain of a heterodimeric receptor molecule, and preferably is an immunoglobulin variable region polypeptide. In a particularly preferred polypeptide V is a V_H or V_L polypeptide. Most preferred is a polypeptide where V is an immunoglobulin V_H polypeptide, and m and n are both zero.

In another embodiment, U or X can define a proteolytic cleavage site, such as the sequence of amino acids found in a precursor protein, such as prothrombin, factor X and the like, that defines the site of cleavage of the protein. A fusion protein having a cleavage site provides a means to purify the protein away from the phage particle to which it is attached.

The polypeptide spacers U and X can each have any sequence of amino acid residues of from about 1 to 6 amino

acid residues in length. Typically the spacer residues are present in a polypeptide to accommodate the continuous reading frame that is required when a polypeptide is produced by the methods disclosed herein using a DNA expression vector of this invention.

A receptor of the present invention assumes a conformation having a binding site specific for, as evidenced by its ability to be competitively inhibited, a preselected or predetermined ligand such as an antigen, enzymatic substrate and the like. In one embodiment, a receptor of this invention is a ligand binding polypeptide that forms an antigen binding site which specifically binds to a preselected antigen to form a complex having a sufficiently strong binding between the antigen and the binding site for the complex to be isolated. When the receptor is an antigen binding polypeptide its affinity or avidity is generally greater than 10^5 M^{-1} more usually greater than 10^6 and preferably greater than 10^8 M^{-1}.

In another embodiment, a receptor of the subject invention binds a substrate and catalyzes the formation of a product from the substrate. While the topology of the ligand binding site of a catalytic receptor is probably more important for its preselected activity than its affinity (association constant or pKa) for the substrate, the subject catalytic receptors have an association constant for the preselected substrate generally greater than 10^3 M^{-1}, more usually greater than 10^5 M^{-1} or 10^6 M^{-1} and preferably greater than 10^7 M^{-1}.

Preferably the receptor produced by the subject invention is heterodimeric and is therefore normally comprised of two different polypeptide chains, which together assume a conformation having a binding affinity, or association constant for the preselected ligand that is different, preferably higher, than the affinity or association constant of either of the polypeptides alone, i.e., as monomers. One or both of the different polypeptide chains is derived from the variable region of the light and heavy chains of an immunoglobulin. Typically, polypeptides comprising the light (V_L) and heavy (V_H) variable regions are employed together for binding the preselected ligand.

A receptor produced by the subject invention can be active in monomeric as well as multimeric forms, either homomeric or heteromeric, preferably heterodimeric. For example, V_H and V_L ligand binding polypeptide produced by the present invention can be advantageously combined in the heterodimer to modulate the activity of either or to produce an activity unique to the heterodimer.

The individual ligand polypeptides will be referred to as V_H and V_L and the heterodimer will be referred to as a Fv. However, it should be understood that a V_H may contain in addition to the V_H, substantially all or a portion of the heavy chain constant region. Similarly, a V_L may contain, in addition to the V_L, substantially all or a portion of the light chain constant region. A heterodimer comprised of a V_H containing a portion of the heavy chain constant region and a V_L containing substantially all of the light chain constant region is termed a Fab fragment. The production of Fab can be advantageous in some situations because the additional constant region sequences contained in a Fab as compared to a Fv can stabilize the V_H and V_L interaction. Such stabilization can cause the Fab to have higher affinity for antigen. In addition the Fab is more commonly used in the art and thus there are more commercial antibodies available to specifically recognize a Fab in screening procedures.

The individual V_H and V_L polypeptides can be produced in lengths equal to or substantially equal to their naturally occurring lengths. However, in preferred embodiments, the V_H and V_L polypeptides will generally have fewer than 125 amino acid residues, more usually fewer than about 120 amino acid residues, while normally having greater than 60 amino acid residues, usually greater than about 95 amino acid residues, more usually greater than about 100 amino acid residues. Preferably, the V_H will be from about 110 to about 230 amino acid residues in length while V_L will be from about 95 to about 214 amino acid residues in length. V_H and V_L chains sufficiently long to form Fabs are preferred.

The amino acid residue sequences will vary widely, depending upon the particular idiotype involved. Usually, there will be at least two cysteines separated by from about 60 to 75 amino acid residues and joined by a disulfide bond. The polypeptides produced by the subject invention will normally be substantial copies of idiotypes of the variable regions of the heavy and/or light chains of immunoglobulins, but in some situations a polypeptide may contain random mutations in amino acid residue sequences in order to advantageously improve the desired activity.

In some situations, it is desirable to provide for covalent cross linking of the V_H and V_L polypeptides, which can be accomplished by providing cysteine resides at the carboxyl termini. The polypeptide will normally be prepared free of the immunoglobulin constant regions, however a small portion of the J region may be included as a result of the advantageous selection of DNA synthesis primers. The D region will normally be included in the transcript of the V_H.

Typically the C terminus region of the V_H and V_L polypeptides will have a greater variety of sequences than the N terminus and, based on the present strategy, can be further modified to permit a variation of the normally occurring V_H and V_L chains. A synthetic polynucleotide can be employed to vary one or more amino acid in a hypervariable region.

In another embodiment, the invention contemplates a heterodimeric receptor molecule comprised of two polypeptide chains, at least one of which is a polypeptide of this invention. Preferably, the polypeptide comprises a receptor domain derived from an immunoglobulin variable chain, more preferably a V_H. More preferred is a heterodimeric receptor comprising receptor domains from both V_H and V_L chains.

E. Methods for Producing a Library

1. General Rationale

In one embodiment the present invention provides a system for the simultaneous cloning and screening of preselected ligand-binding specificities from gene repertoires using a single vector system. This system provides linkage of cloning and screening methodologies and has two requirements. First, that expression of the polypeptide chains of a heterodimeric receptor in an in vitro expression host such as *E. coli* requires coexpression of the two polypeptide chains in order that a functional heterodimeric receptor can assemble to produce a receptor that binds ligand. Second, that screening of isolated members of the library for a preselected ligand-binding capacity requires a means to correlate the binding capacity of an expressed receptor molecule with a convenient means to isolate the gene that encodes the member from the library.

Linkage of expression and screening is accomplished by the combination of targeting of a fusion protein into the periplasm of a bacterial cell to allow assembly of a functional receptor, and the targeting of a fusion protein onto the coat of a filamentous phage particle during phage assembly to allow for convenient screening of the library member of interest. Periplasmic targeting is provided by the presence of a secretion signal domain in a fusion protein of this inven-

tion. Targeting to a phage particle is provided by the presence of a filamentous phage coat protein membrane anchor domain in a fusion protein of this invention.

The present invention describes in one embodiment a method for producing a library of DNA molecules, each DNA molecule comprising a cistron for expressing a fusion protein on the surface of a filamentous phage particle. The method comprises the steps of (a) forming a ligation admixture by combining in a ligation buffer (i) a repertoire of polypeptide encoding genes and (ii) a plurality of DNA expression vectors in linear form adapted to form a fusion protein expressing cistron, and (b) subjecting the admixture to ligation conditions for a time period sufficient for the repertoire of genes to become operatively linked (ligated) to the plurality of vectors to form the library.

In this embodiment, the repertoire of polypeptide encoding genes are in the form of double-stranded (ds) DNA and each member of the repertoire has cohesive termini adapted for directional ligation. In addition, the plurality of DNA expression vectors are each linear DNA molecules having upstream and downstream cohesive termini that are (a) adapted for directionally receiving the polypeptide genes in a common reading frame, and (b) operatively linked to respective upstream and downstream translatable DNA sequences. The upstream translatable DNA sequence encodes a secretion signal, preferably a pelB secretion signal, and the downstream translatable DNA sequence encodes a filamentous phage coat protein membrane anchor as described herein for a polypeptide of this invention. The translatable DNA sequences are also operatively linked to respective upstream and downstream DNA expression control sequences as defined for a DNA expression vector described herein.

The library so produced can be utilized for expression and screening of the fusion proteins encoded by the resulting library of cistrons represented in the library by the expression and screening methods described herein.

2. Production of Gene Repertoires

A gene repertoire is a collection of different genes, preferably polypeptide-encoding genes (polypeptide genes), and may be isolated from natural sources or can be generated artificially. Preferred gene repertoires are comprised of conserved genes. Particularly preferred gene repertoires comprise either or both genes that code for the members of a dimeric receptor molecule.

A gene repertoire useful in practicing the present invention contains at least 10^3, preferably at least 10^4, more preferably at least 10^5, and most preferably at least 10^7 different genes. Methods for evaluating the diversity of a repertoire of genes is well known to one skilled in the art.

Thus, in one embodiment, the present invention contemplates a method of isolating a pair of genes coding for a dimeric receptor having a preselected activity from a repertoire of conserved genes. Additionally, expressing the cloned pair of genes and isolating the resulting expressed dimeric receptor protein is also described. Preferably, the receptor will be a heterodimeric polypeptide capable of binding a ligand, such as an antibody molecule or immunologically active portion thereof, a cellular receptor, or a cellular adhesion protein coded for by one of the members of a family of conserved genes, i.e., genes containing a conserved nucleotide sequence of at least about 10 nucleotides in length.

Exemplary conserved gene families encoding different polypeptide chains of a dimeric receptor are those coding for immunoglobulins, major histocompatibility complex antigens of class I or II, lymphocyte receptors, integrins and the like.

A gene can be identified as belonging to a repertoire of conserved genes using several methods. For example, an isolated gene may be used as a hybridization probe under low stringency conditions to detect other members of the repertoire of conserved genes present in genomic DNA using the methods described by Southern, *J. Mol. Biol.*, 98:503 (1975). If the gene used as a hybridization probe hybridizes to multiple restriction endonuclease fragments of the genome, that gene is a member of a repertoire of conserved genes.

Immunoglobulins

The immunoglobulins, or antibody molecules, are a large family of molecules that include several types of molecules, such as IgD, IgG, IgA, IgM and IgE. The antibody molecule is typically comprised of two heavy (H) and light (L) chains with both a variable (V) and constant (C) region present on each chain as shown in FIG. 1. Schematic diagrams of human IgG heavy chain and human kappa light chain are shown in FIGS. 2A and 2B, respectively. Several different regions of an immunoglobulin contain conserved sequences useful for isolating an immunoglobulin repertoire. Extensive amino acid and nucleic acid sequence data displaying exemplary conserved sequences is compiled for immunoglobulin molecules by Kabat et al., in *Sequences of Proteins of Immunological Interest*, National Institutes of Health, Bethesda, Md., 1987.

The C region of the H chain defines the particular immunoglobulin type. Therefore the selection of conserved sequences as defined herein from the C region of the H chain results in the preparation of a repertoire of immunoglobulin genes having members of the immunoglobulin type of the selected C region.

The V region of the H or L chain typically comprises four framework (FR) regions each containing relatively lower degrees of variability that includes lengths of conserved sequences. The use of conserved sequences from the FR1 and FR4 (J region) framework regions of the V_H chain is a preferred exemplary embodiment and is described herein in the Examples. Framework regions are typically conserved across several or all immunoglobulin types and thus conserved sequences contained therein are particularly suited for preparing repertoires having several immunoglobulin types.

Major Histocompatibility Complex

The major histocompatibility complex (MHC) is a large genetic locus that encodes an extensive family of proteins that include several classes of molecules referred to as class I, class II or class III MHC molecules. Paul et al., in *Fundamental Immunology*, Raven Press, N.Y., pp. 303–378 (1984).

Class I MHC molecules are a polymorphic group of transplantation antigens representing a conserved family in which the antigen is comprised of a heavy chain and a non-MHC encoded light chain. The heavy chain includes several regions, termed the N, C1, C2, membrane and cytoplasmic regions. Conserved sequences useful in the present invention are found primarily in the N, C1 and C2 regions and are identified as continuous sequences of "invariant residues" in Kabat et al., supra.

Class II MHC molecules comprise a conserved family of polymorphic antigens that participate in immune responsiveness and are comprised of an alpha and a beta chain. The genes coding for the alpha and beta chain each include several regions that contain conserved sequences suitable for producing MHC class II alpha or beta chain repertoires. Exemplary conserved nucleotide sequences include those coding for amino acid residues 26–30 of the A1 region,

residues 161–170 of the A2 region and residues 195–206 of the membrane region, all of the alpha chain. Conserved sequences are also present in the B1, B2 and membrane regions of the beta chain at nucleotide sequences coding for amino acid residues 41–45, 150–162 and 200–209, respectively.

Lymphocyte Receptors and Cell Surface Antigens

Lymphocytes contain several families of proteins on their cell surfaces including the T-cell receptor, Thy-1 antigen and numerous T-cell surface antigens including the antigens defined by the monoclonal antibodies OKT4 (leu3), OKT5/8 (leu2), OKT3, OKT1 (leu1), OKT 11 (leu5) OKT6 and OKT9. Paul, supra at pp. 458–479.

The T-cell receptor is a term used for a family of antigen binding molecules found on the surface of T-cells. The T-cell receptor as a family exhibits polymorphic binding specificity similar to immunoglobulins in its diversity. The mature T-cell receptor is comprised of alpha and beta chains each having a variable (V) and constant (C) region. The similarities that the T-cell receptor has to immunoglobulins in genetic organization and function shows that T-cell receptor contains regions of conserved sequence. Lai et al., *Nature*, 331:543–546 (1988).

Exemplary conserved sequences include those coding for amino acid residues 84–90 of alpha chain, amino acid residues 107–115 of beta chain, and amino acid residues 91–95 and 111–116 of the gamma chain. Kabat et al., supra, p. 279.

Integrins And Adhesions

Adhesive proteins involved in cell attachment are members of a large family of related proteins termed integrins. Integrins are heterodimers comprised of a beta and an alpha subunit. Members of the integrin family include the cell surface glycoproteins platelet receptor GpIIb-IIIa, vitronectin receptor (VnR), fibronectin receptor (FnR) and the leukocyte adhesion receptors LFA-1, Mac-1, Mo-1 and 60.3. Rouslahti et al., *Science*, 238:491–497 (1987). Nucleic acid and protein sequence data demonstrates regions of conserved sequences exist in the members of these families, particularly between the beta chain of GpIIb-IIIa, VnR and FnR, and between the alpha subunit of VnR, Mac-1, LFA-1, FnR and GpIIb-IIIa. Suzuki et al., *Proc. Natl. Acad. Sci. USA*, 83:8614–8618, 1986; Ginsberg et al., *J. Biol. Chem.*, 262:5437–5440, 1987.

Various well known methods can be employed to produce a useful gene repertoire. For instance, V_H and V_L gene repertoires can be produced by isolating V_H- and V_L-coding mRNA from a heterogeneous population of antibody producing cells, i.e., B lymphocytes (B cells), preferably rearranged B cells such as those found in the circulation or spleen of a vertebrate. Rearranged B cells are those in which immunoglobulin gene translocation, i.e., rearrangement, has occurred as evidenced by the presence in the cell of mRNA with the immunoglobulin gene V, D and J region transcripts adjacently located thereon. Typically, the B cells are collected in a 1–100 ml sample of blood which usually contains 10^6 B cells/ml.

In some cases, it is desirable to bias a repertoire for a preselected activity, such as by using as a source of nucleic acid cells (source cells) from vertebrates in any one of various stages of age, health and immune response. For example, repeated immunization of a healthy animal prior to collecting rearranged B cells results in obtaining a repertoire enriched for genetic material producing a receptor of high affinity. Mullinax et al., *Proc. Natl. Acad. Sci. USA*, 87:8095–8099 (1990). Conversely, collecting rearranged B cells from a healthy animal whose immune system has not been recently challenged results in producing a repertoire that is not biased towards the production of high affinity V_H and/or V_L polypeptides.

It should be noted the greater the genetic heterogeneity of the population of cells for which the nucleic acids are obtained, the greater the diversity of the immunological repertoire (comprising V_H- and V_L-coding genes) that will be made available for screening according to the method of the present invention. Thus, cells from different individuals, particularly those having an immunologically significant age difference, and cells from individuals of different strains, races or species can be advantageously combined to increase the heterogeneity (diversity) of a repertoire.

Thus, in one preferred embodiment, the source cells are obtained from a vertebrate, preferably a mammal, which has been immunized or partially immunized with an antigenic ligand (antigen) against which activity is sought, i.e., a preselected antigen. The immunization can be carried out conventionally. Antibody titer in the animal can be monitored to determine the stage of immunization desired, which stage corresponds to the amount of enrichment or biasing of the repertoire desired. Partially immunized animals typically receive only one immunization and cells are collected from those animals shortly after a response is detected. Fully immunized animals display a peak titer, which is achieved with one or more repeated injections of the antigen into the host mammal, normally at 2 to 3 week intervals. Usually three to five days after the last challenge, the spleen is removed and the genetic repertoire of the splenocytes, about 90% of which are rearranged B cells, is isolated using standard procedures. See, *Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley & Sons, NY. Nucleic acids coding for V_H and V_L polypeptides can be derived from cells producing IgA, IgD, IgE, IgG or IgM, most preferably from IgM and IgG, producing cells.

Methods for preparing fragments of genomic DNA from which immunoglobulin variable region genes can be cloned as a diverse population are well known in the art. See for example Herrmann et al., *Methods In Enzymol.*, 152:180–183, (1987); Frischauf, *Methods In Enzymol.*, 152:183–190 (1987); Frischauf, *Methods In Enzymol.*, 152:190–199 (1987); and DiLella et al., *Methods In Enzymol.*, 152:199–212 (1987). (The teachings of the references cited herein are hereby incorporated by reference.)

The desired gene repertoire can be isolated from either genomic material containing the gene expressing the variable region or the messenger RNA (mRNA) which represents a transcript of the variable region. The difficulty in using the genomic DNA from other than non-rearranged B lymphocytes is in juxtaposing the sequences coding for the variable region, where the sequences are separated by introns. The DNA fragment(s) containing the proper exons must be isolated, the introns excised, and the exons then spliced in the proper order and in the proper orientation. For the most part, this will be difficult, so that the alternative technique employing rearranged B cells will be the method of choice because the V, D and J immunoglobulin gene regions have translocated to become adjacent, so that the sequence is continuous (free of introns) for the entire variable regions.

Where mRNA is utilized the cells will be lysed under RNase inhibiting conditions. In one embodiment, the first step is to isolate the total cellular mRNA. Poly A+ mRNA can then be selected by hybridization to an oligo-dT cellulose column. The presence of mRNAs coding for the heavy and/or light chain polypeptides can then be assayed by hybridization with DNA single strands of the appropriate

genes. Conveniently, the sequences coding for the constant portion of the V_H and V_L can be used as polynucleotide probes, which sequences can be obtained from available sources. See for example, Early and Hood, *Genetic Engineering*, Setlow and Hollaender, eds., Vol. 3, Plenum Publishing Corporation, NY, (1981), pages 157–188; and Kabat et al., *Sequences of Immunological Interest*, National Institutes of Health, Bethesda, Md., (1987).

In preferred embodiments, the preparation containing the total cellular mRNA is first enriched for the presence of V_H and/or V_L coding mRNA. Enrichment is typically accomplished by subjecting the total mRNA preparation or partially purified mRNA product thereof to a primer extension reaction employing a polynucleotide synthesis primer as described herein. Exemplary methods for producing V_H and V_L gene repertoires using polynucleotide synthesis primers are described in PCT Application No. PCT/US 90/02836 (International Publication No. WO 90/14430). Particularly preferred methods for producing a gene repertoire rely on the use of preselected oligonucleotides as primers in a polymerase chain reaction (PCR) to form PCR reaction products as described herein.

In preferred embodiments, isolated B cells are immunized in vitro against a preselected antigen. In vitro immunization is defined as the clonal expansion of epitope-specific B cells in culture, in response to antigen stimulation. The end result is to increase the frequency of antigen-specific B cells in the immunoglobulin repertoire, and thereby decrease the number of clones in an expression library that must be screened to identify a clone expressing an antibody of the desired specificity. The advantage of in vitro immunization is that human monoclonal antibodies can be generated against a limitless number of therapeutically valuable antigens, including toxic or weak immunogens. For example, antibodies specific for the polymorphic determinants of tumor-associated antigens, rheumatoid factors, and histocompatibility antigens can be produced, which can not be elicited in immunized animals. In addition, it may be possible to generate immune responses which are normally suppressed in vivo.

In vitro immunization can be used to give rise to either a primary or secondary immune response. A primary immune response, resulting from first time exposure of a B cell to an antigen, results in clonal expansion of epitope-specific cells and the secretion of IgM antibodies with low to moderate apparent affinity constants (10^6–10^8 M^{-1}). Primary immunization of human splenic and tonsillar lymphocytes in culture can be used to produce monoclonal antibodies against a variety of antigens, including cells, peptides, macromolecule, haptens, and tumor-associated antigens. Memory B cells from immunized donors can also be stimulated in culture to give rise to a secondary immune response characterized by clonal expansion and the production of high affinity antibodies (>10^9 M^{-1}) of the IgG isotype, particularly against viral antigens by clonally expanding sensitized lymphocytes derived from seropositive individuals.

In one embodiment, peripheral blood lymphocytes are depleted of various cytolytic cells that appear to down-modulate antigen-specific B cell activation. When lysosome-rich subpopulations (natural killer cells, cytotoxic and suppressor T cells, monocytes) are first removed by treatment with the lysosmotropic methyl ester of leucine, the remaining cells (including B cells, T helper cells, accessory cells) respond antigen-specifically during in vitro immunization. The lymphokine requirements for inducing antibody production in culture are satisfied by a culture supernatant from activated, irradiated T cells.

In addition to in vitro immunization, cell panning (immunoaffinity absorption) can be used to further increase the frequency of antigen-specific B cells. Techniques for selecting B cell subpopulations via solid-phase antigen binding are well established. Panning conditions can be optimized to selectively enrich for B cells which bind with high affinity to a variety of antigens, including cell surface proteins. Panning can be used alone, or in combination with in vitro immunization to increase the frequency of antigen-specific cells above the levels which can be obtained with either technique alone. Immunoglobulin expression libraries constructed from enriched populations of B cells are biased in favor of antigen-specific antibody clones, and thus, enabling identification of clones with the desired specificities from smaller, less complex libraries.

3. Preparation of Polynucleotide Primers

The term "polynucleotide" as used herein in reference to primers, probes and nucleic acid fragments or segments to be synthesized by primer extension is defined as a molecule comprised of two or more deoxyribonucleotide or ribonucleotides, preferably more than 3. Its exact size will depend on many factors, which in turn depends on the ultimate conditions of use.

The term "primer" as used herein refers to a polynucleotide whether purified from a nucleic acid restriction digest or produced synthetically, which is capable of acting as a point of initiation of nucleic acid synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase, reverse transcriptase and the like, and at a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency, but may alternatively be in double stranded form. If double stranded, the primer is first treated to separate it from its complementary strand before being used to prepare extension products. Preferably, the primer is a polydeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agents for polymerization. The exact lengths of the primers will depend on may factors, including temperature and the source of primer. For example, depending on the complexity of the target sequence, a polynucleotide primer typically contains 15 to 25 or more nucleotides, although it can contain fewer nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with template.

The primers used herein are selected to be "substantially" complementary to the different strands of each specific sequence to be synthesized or amplified. This means that the primer must be sufficiently complementary to non-randomly hybridize with its respective template strand. Therefore, the primer sequence may or may not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment can be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Such non-complementary fragments typically code for an endonuclease restriction site. Alternatively, non-complementary bases or longer sequences can be interspersed into the primer, provided the primer sequence has sufficient complementarily with the sequence of the strand to be synthesized or amplified to non-randomly hybridize therewith and thereby form an extension product under polynucleotide synthesizing conditions.

Primers of the present invention may also contain a DNA-dependent RNA polymerase promoter sequence or its

complement. See for example, Krieg et al., *Nucl. Acids Res.*, 12:7057–70 (1984); Studier et al., *J. Mol. Biol.*, 189:113–130 (1986); and *Molecular Cloning: A Laboratory Manual, Second Edition*, Maniatis et al., eds., Cold Spring Harbor, N.Y. (1989).

When a primer containing a DNA-dependent RNA polymerase promoter is used the primer is hybridized to the polynucleotide strand to be amplified and the second polynucleotide strand of the DNA-dependent RNA polymerase promoter is completed using an inducing agent such as *E. coli* DNA polymerase I, or the Klenow fragment of *E. coli* DNA polymerase. The starting polynucleotide is amplified by alternating between the production of an RNA polynucleotide and DNA polynucleotide.

Primers may also contain a template sequence or replication initiation site for a RNA-directed RNA polymerase. Typical RNA-directed RNA polymerase include the QB replicase described by Lizardi et al., *Biotechnology*, 6:1197–1202 (1988) RNA-directed polymerases produce large numbers of RNA strands from a small number of template RNA strands that contain a template sequence or replication initiation site. These polymerases typically give a one million-fold amplification of the template strand as has been described by Kramer et al., *J. Mol. Biol.*, 89:719–736 (1974).

The polynucleotide primers can be prepared using any suitable method, such as, for example, the phosphotriester or phosphodiester methods see Narang et al., *Meth. Enzymol.*, 68:90, (1979); U.S. Pat. No. 4,356,270; and Brown et al., *Meth. Enzymol.*, 68:109, (1979).

The choice of a primer's nucleotide sequence depends on factors such as the distance on the nucleic acid from the region coding for the desired receptor, its hybridization site on the nucleic acid relative to any second primer to be used, the number of genes in the repertoire it is to hybridize to, and the like.

a. Primers for Producing Immunoglobulin Gene Repertoires

V_H and V_L gene repertoires can be separately prepared prior to their utilization in the present invention. Repertoire preparation is typically accomplished by primer extension, preferably by primer extension in a polymerase chain reaction (PCR) format.

To produce a repertoire of V_H-coding DNA homologs by primer extension, the nucleotide sequence of a primer is selected to hybridize with a plurality of immunoglobulin heavy chain genes at a site substantially adjacent to the V_H-coding region so that a nucleotide sequence coding for a functional (capable of binding) polypeptide is obtained. To hybridize to a plurality of different V_H-coding nucleic acid strands, the primer must be a substantial complement of a nucleotide sequence conserved among the different strands. Such sites include nucleotide sequences in the constant region, any of the variable region framework regions, preferably the third framework region, leader region, promoter region, J region and the like.

If the repertoires of V_H-coding and V_L-coding DNA homologs are to be produced by (PCR) amplification, two primers, i.e., a PCR primer pair, must be used for each coding strand of nucleic acid to be amplified. The first primer becomes part of the nonsense (minus or complementary) strand and hybridizes to a nucleotide sequence conserved among V_H (plus or coding) strands within the repertoire. To produce V_H coding DNA homologs, first primers are therefore chosen to hybridize to (i.e. be complementary to) conserved regions within the J region, CH1 region, hinge region, CH2 region, or CH3 region of immunoglobulin genes and the like. To produce a V_L coding DNA homolog, first primers are chosen to hybridize with (i.e. be complementary to) a conserved region within the J region or constant region of immunoglobulin light chain genes and the like. Second primers become part of the coding (plus) strand and hybridize to a nucleotide sequence conserved among minus strands. To produce the V_H-coding DNA homologs, second primers are therefore chosen to hybridize with a conserved nucleotide sequence at the 5' end of the V_H-coding immunoglobulin gene such as in that area coding for the leader or first framework region. It should be noted that in the amplification of both V_H- and V_L-coding DNA homologs the conserved 5' nucleotide sequence of the second primer can be complementary to a sequence exogenously added using terminal deoxynucleotidyl transferase as described by Loh et al., *Science*, 243:217–220 (1989). One or both of the first and second primers can contain a nucleotide sequence defining an endonuclease recognition site. The site can be heterologous to the immunoglobulin gene being amplified and typically appears at or near the 5' end of the primer.

When present, the restriction site-defining portion is typically located in a 5'-terminal non-priming portion of the primer. The restriction site defined by the first primer is typically chosen to be one recognized by a restriction enzyme that does not recognize the restriction site defined by the second primer, the objective being to be able to produce a DNA molecule having cohesive termini that are non-complementary to each other and thus allow directional insertion into a vector.

In one embodiment, the present invention utilizes a set of polynucleotides that form primers having a priming region located at the 3'-terminus of the primer. The priming region is typically the 3'-most (3'-terminal) 15 to 30 nucleotide bases. The 3'-terminal priming portion of each primer is capable of acting as a primer to catalyze nucleic acid synthesis, i.e., initiate a primer extension reaction off its 3' terminus. One or both of the primers can additionally contain a 5'-terminal (5'-most) non-priming portion, i.e., a region that does not participate in hybridization to repertoire template.

In PCR, each primer works in combination with a second primer to amplify a target nucleic acid sequence. The choice of PCR primer pairs for use in PCR is governed by considerations as discussed herein for producing gene repertoires. That is, the primers have a nucleotide sequence that is complementary to a sequence conserved in the repertoire. Useful V_H and V_L priming sequences are shown in Tables 5 and 6, herein below.

4. Polymerase Chain Reaction to Produce Gene Repertoires

The strategy used for cloning the V_H and V_L genes contained within a repertoire will depend, as is well known in the art, on the type, complexity, and purity of the nucleic acids making up the repertoire. Other factors include whether or not the genes are contained in one or a plurality of repertoires and whether or not they are to be amplified and/or mutagenized.

The V_H- and V_L-coding gene repertoires are comprised of polynucleotide coding strands, such as mRNA and/or the sense strand of genomic DNA. If the repertoire is in the form of double stranded genomic DNA, it is usually first denatured, typically by melting, into single strands. A repertoire is subjected to a PCR reaction by treating (contacting) the repertoire with a PCR primer pair, each member of the pair having a preselected nucleotide sequence. The PCR primer pair is capable of initiating primer extension reactions by hybridizing to nucleotide sequences, preferably at least about 10 nucleotides in length

and more preferably at least about 20 nucleotides in length, conserved within the repertoire. The first primer of a PCR primer pair is sometimes referred to herein as the "sense primer" because it hybridizes to the coding or sense strand of a nucleic acid. In addition, the second primer of a PCR primer pair is sometimes referred to herein as the "anti-sense primer" because it hybridizes to a non-coding or anti-sense strand of a nucleic acid, i.e., a strand complementary to a coding strand.

The PCR reaction is performed by mixing the PCR primer pair, preferably a predetermined amount thereof, with the nucleic acids of the repertoire, preferably a predetermined amount thereof, in a PCR buffer to form a PCR reaction admixture. The admixture is maintained under polynucleotide synthesizing conditions for a time period, which is typically predetermined, sufficient for the formation of a PCR reaction product, thereby producing a plurality of different V_H-coding and/or V_L-coding DNA homologs.

A plurality of first primer and/or a plurality of second primers can be used in each amplification, e.g., one species of first primer can be paired with a number of different second primers to form several different primer pairs. Alternatively, an individual pair of first and second primers can be used. In any case, the amplification products of amplifications using the same or different combinations of first and second primers can be combined to increase the diversity of the gene library.

In another strategy, the object is to clone the V_H- and/or V_L-coding genes from a repertoire by providing a polynucleotide complement of the repertoire, such as the anti-sense strand of genomic dsDNA or the polynucleotide produced by subjecting mRNA to a reverse transcriptase reaction. Methods for producing such complements are well known in the art.

The PCR reaction is performed using any suitable method. Generally it occurs in a buffered aqueous solution, i.e., a PCR buffer, preferably at a pH of 7–9, most preferably about 8. Preferably, a molar excess (for genomic nucleic acid, usually about 10^6:1 primer:template) of the primer is admixed to the buffer containing the template strand. A large molar excess is preferred to improve the efficiency of the process.

The PCR buffer also contains the deoxyribonucleotide triphosphates dATP, dCTP, dGTP, and dTTP and a polymerase, typically thermostable, all in adequate amounts for primer extension (polynucleotide synthesis) reaction. The resulting solution (PCR admixture) is heated to about 90° C. -100° C. for about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period the solution is allowed to cool to 54° C., which is preferable for primer hybridization. The synthesis reaction may occur at from room temperature up to a temperature above which the polymerase (inducing agent) no longer functions efficiently. Thus, for example, if DNA polymerase is used as inducing agent, the temperature is generally no greater than about 40° C. An exemplary PCR buffer comprises the following: 50 mM KCl; 10 mM Tris-HCl; pH 8.3; 1.5 mM $MgCl_2$; 0.001% (wt/vol) gelatin, 200 μM dATP; 200 μM dTTP; 200 μM dCTP; 200 μM dGTP; and 2.5 units *Thermus aquaticus* DNA polymerase I (U.S. Pat. No. 4,889,818) per 100 microliters of buffer.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase I, T4 DNA polymerase, other available DNA polymerases, reverse transcriptase, and other enzymes, including heat-stable enzymes, which will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths. There may be inducing agents, however, which initiate synthesis at the 5' end and proceed in the above direction, using the same process as described above.

The inducing agent also may be a compound or system which will function to accomplish the synthesis of RNA primer extension products, including enzymes. In preferred embodiments, the inducing agent may be a DNA-dependent RNA polymerase such as T7 RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase. These polymerases produce a complementary RNA polynucleotide. The high turn over rate of the RNA polymerase amplifies the starting polynucleotide as has been described by Chamberlin et al., *The Enzymes*, ed. P. Boyer, PP. 87–108, Academic Press, New York (1982). Another advantage of T7 RNA polymerase is that mutations can be introduced into the polynucleotide synthesis by replacing a portion of cDNA with one or more mutagenic oligodeoxynucleotides (polynucleotides) and transcribing the partially-mismatched template directly as has been previously described by Joyce et al., *Nuc. Acid Res.*, 17:711–722 (1989). Amplification systems based on transcription have been described by Gingeras et al., in *PCR Protocols, A Guide to Methods and Applications*, pp 245–252, Academic Press, Inc., San Diego, Calif. (1990).

If the inducing agent is a DNA-dependent RNA polymerase and therefore incorporates ribonucleotide triphosphates, sufficient amounts of ATP, CTP, GTP and UTP are admixed to the primer extension reaction admixture and the resulting solution is treated as described above.

The newly synthesized strand and its complementary nucleic acid strand form a double-stranded molecule which can be used in the succeeding steps of the process.

The first and/or second PCR reactions discussed above can advantageously be used to incorporate into the receptor a preselected epitope useful in immunologically detecting and/or isolating a receptor. This is accomplished by utilizing a first and/or second polynucleotide synthesis primer or expression vector to incorporate a predetermined amino acid residue sequence into the amino acid residue sequence of the receptor.

After producing V_H- and V_L-coding DNA homologs for a plurality of different V_H- and V_L-coding genes within the repertoires, the DNA molecules are typically further amplified. While the DNA molecules can be amplified by classic techniques such as incorporation into an autonomously replicating vector, it is preferred to first amplify the molecules by subjecting them to a polymerase chain reaction (PCR) prior to inserting them into a vector. PCR is typically carried out by thermocycling i.e., repeatedly increasing and decreasing the temperature of a PCR reaction admixture within a temperature range whose lower limit is about 10° C. to about 40° C. and whose upper limit is about 90° C. to about 100° C. The increasing and decreasing can be continuous, but is preferably phasic with time periods of relative temperature stability at each of temperatures favoring polynucleotide synthesis, denaturation and hybridization.

PCR amplification methods are described in detail in U.S. Pat. Nos. 4,683,192, 4,683,202, 4,800,159, and 4,965,188,

and at least in several texts including "PCR Technology: Principles and Applications for DNA Amplification". H. Erlich. ed.. Stockton Press. New York (1989); and "PCR Protocols: A Guide to Methods and Applications". Innis et al.. eds.. Academic Press. San Diego. Calif. (1990).

In preferred embodiments only one pair of first and second primers is used per amplification reaction. The amplification reaction products obtained from a plurality of different amplifications. each using a plurality of different primer pairs. are then combined.

However. the present invention also contemplates DNA homolog production via co-amplification (using two pairs of primers). and multiplex amplification (using up to about 8. 9 or 10 primer pairs).

In preferred embodiments. the PCR process is used not only to produce a library of DNA molecules. but also to induce mutations within the library or to create diversity from a single parental clone and thereby provide a library having a greater heterogeneity. First. it should be noted that the PCR process itself is inherently mutagenic due to a variety of factors well known in the art. Second. in addition to the mutation inducing variations described in the above referenced U.S. Pat. No. 4.683.195. other mutation inducing PCR variations can be employed. For example. the PCR reaction admixture. can be formed with different amounts of one or more of the nucleotides to be incorporated into the extension product. Under such conditions. the PCR reaction proceeds to produce nucleotide substitutions within the extension product as a result of the scarcity of a particular base. Similarly. approximately equal molar amounts of the nucleotides can be incorporated into the initial PCR reaction admixture in an amount to efficiently perform X number of cycles. and then cycling the admixture through a number of cycles in excess of X. such as. for instance. 2X. Alternatively. mutations can be induced during the PCR reaction by incorporating into the reaction admixture nucleotide derivatives such as inosine. not normally found in the nucleic acids of the repertoire being amplified. During subsequent in vivo amplification. the nucleotide derivative will be replaced with a substitute nucleotide thereby inducing a point mutation.

5. Linear DNA Expression Vectors

A DNA expression vector for use in a method of the invention for producing a library of DNA molecules is a linearized DNA molecule as described before having two (upstream and downstream) cohesive termini adapted for directional ligation to a polypeptide gene.

A linear DNA expression vector is typically prepared by restriction endonuclease digestion of a circular DNA expression vector of this invention to cut at two preselected restriction sites within the sequence of nucleotides of the vector adapted for directional ligation to produce a linear DNA molecule having the required cohesive termini that are adapted for direction ligation. Directional ligation refers to the presence of two (a first and second) cohesive termini on a vector. or on the insert DNA molecule to be ligated into the vector selected. so that the termini on a single molecule are not complementary. A first terminus of the vector is complementary to a first terminus of the insert. and the second terminus of the vector is complementary to the second terminus of the insert.

6. Ligation Reactions to Produce Gene Libraries

In preparing a library of DNA molecules of this invention. a ligation admixture is prepared as described above. and the admixture is subjected to ligation conditions for a time period sufficient for the admixed repertoire of polypeptide genes to ligate (become operatively linked) to the plurality of DNA expression vectors to form the library.

Ligation conditions are conditions selected to favor a ligation reaction wherein a phosphodiester bond is formed between adjacent 3' hydroxyl and 5' phosphoryl termini of DNA. The ligation reaction is preferably catalyzed by the enzyme T4 DNA ligase. Ligation conditions can vary in time. temperature. concentration of buffers. quantities of DNA molecules to be ligated. and amounts of ligase. as is well known. Preferred ligation conditions involve maintaining the ligation admixture at 4 degrees Centigrade (4° C.) to 12° C. for 1 to 24 hours in the presence of 1 to 10 units of T4 DNA ligase per milliliter (ml) and about 1 to 2 micrograms (ug) of DNA. Ligation buffer in a ligation admixture typically contains 0.5M Tris-HCl (pH 7.4). 0.01M $MgCl_2$. 0.01M dithiothrietol. 1 mM spermidine. 1 mM ATP and 0.1 mg/ml bovine serum albumin (BSA). Other ligation buffers can also be used.

Exemplary ligation reactions are described in Example 2.

7. Preparation of Dicistronic Gene Libraries

In a particularly preferred embodiment. the present invention contemplates methods for the preparation of a library of dicistronic DNA molecules. A dicistronic DNA molecule is a single DNA molecule having the capacity to express two separate polypeptides from two separate cistrons. In preferred embodiments. the two cistrons are operatively linked at relative locations on the DNA molecule such that both cistrons are under the transcriptional control of a single promoter. Each dicistronic molecule is capable of expressing first and second polypeptides from first and second cistrons. respectively. that can form. in a suitable host. a heterodimeric receptor on the surface of a filamentous phage particle.

The method for producing a library of dicistronic DNA molecules comprises the steps of:

(a) Forming a first ligation admixture by combining in a ligation buffer:

(i) a repertoire of first polypeptide genes in the form of dsDNA. each having cohesive termini adapted for directional ligation. and

(ii) a plurality of DNA expression vectors in linear form. each having upstream and downstream first cohesive termini that are (a) adapted for directionally receiving the first polypeptide genes in a common reading frame. and (b) operatively linked to respective upstream and downstream translatable DNA sequences. The upstream translatable DNA sequence encodes a pelB secretion signal. the downstream translatable DNA sequence encodes a filamentous phage coat protein membrane anchor. and translatable DNA sequences are operatively linked to respective upstream and downstream DNA expression control sequences.

(b) Subjecting the admixture to ligation conditions for a time period sufficient to operatively link the first polypeptide genes to the vectors and produce a plurality of circular DNA molecules each having a first cistron for expressing the first polypeptide.

(c) Treating the plurality of circular DNA molecules under DNA cleavage conditions to produce a plurality of DNA expression vectors in linear form that each have upstream and downstream second cohesive termini that are (i) adapted for directionally receiving a repertoire of second polypeptide genes in a common reading frame. and (ii) operatively linked to respective upstream and downstream DNA sequences. The upstream DNA sequence is a translatable sequence encoding a secretion signal. the downstream DNA sequence has at least one stop codon in the reading frame. and the translat-

able DNA sequence is operatively linked to a DNA expression control sequence.

(d) Forming a second ligation admixture by combining in a ligation buffer:

(i) the plurality of DNA expression vectors formed in step (c), and

(ii) the repertoire of second polypeptide genes in the form of dsDNA, each having cohesive termini adapted for directional ligation to the plurality of DNA expression vectors; and

(e) Subjecting the second admixture to ligation conditions for a time period sufficient to operatively link the second polypeptide genes to said vectors and produce a plurality of circular DNA molecules each having the second cistron for expressing the second polypeptide, thereby forming the library. In preferred embodiments a secretion signal is a pelB secretion signal, and the membrane anchor is derived from cpVIII as described herein.

DNA expression vectors useful for practicing the above method are the dicistronic expression vectors described in greater detail before.

In practicing the method of producing a library of dicistronic DNA molecules, it is preferred that the upstream and downstream first cohesive termini do not have the same nucleotide sequences as the upstream and downstream second cohesive termini. In this embodiment, the treating step (c) to linearize the circular DNA molecules typically involves the use of restriction endonucleases that are specific for producing said second termini, but do not cleave the circular DNA molecule at the sites that formed the first termini. Exemplary and preferred first and second termini are the termini defined by cleavage of pCBAK8 with Xho I and Spe I to form the upstream and downstream first termini, and defined by cleavage of pCBAK8 with Sac I and Xba I to form the upstream and downstream second termini. In this embodiment, other pairs of cohesive termini can be utilized at the respective pairs of first and second termini, so long as the four termini are each distinct, non-complementary termini.

Methods of treating the plurality of circular DNA molecules under DNA cleavage conditions to form linear DNA molecules are generally well known and depend on the nucleotide sequence to be cleaved and the mechanism for cleavage. Preferred treatments involve admixing the DNA molecules with a restriction endonuclease specific for a endonuclease recognition site at the desired cleavage location in an amount sufficient for the restriction endonuclease to cleave the DNA molecule. Buffers, cleavage conditions, and substrate concentrations for restriction endonuclease cleavage are well known and depend on the particular enzyme utilized. Exemplary restriction enzyme cleavage conditions are described in Example 2.

F. Phage Libraries

The present invention contemplates a library of DNA molecules that each encode a fusion protein of this invention where the library is in the form of a population of different filamentous phage particles each containing one of the different rDNA molecules. By different ,DNA molecule is meant rDNA molecules differing in nucleotide base sequence encoding a polypeptide of this invention.

Thus, a phage library is a population of filamentous phage, preferably f1, fd or M13 filamentous phage, each phage having packaged inside the particle a rDNA expression vector of this invention. A preferred library is comprised of phage particles containing DNA molecules that encode at least 10^6, preferably 10^7 and more preferably 10^{8-9} different fusion proteins of this invention. By different fusion proteins is meant fusion proteins differing in amino acid residue sequence. Where the packaged expression vector encodes first and second polypeptides of an autogenously assembling receptor, e.g. V_H and V_L polypeptides that form a Fab, the library can also be characterized as containing or expressing a multiplicity of receptor specificities. Thus, preferred libraries express at least 10^5, preferably at least 10^6 and more preferably at least 10^7 different receptors, such as different antibodies, T cell receptors, integrins and the like.

As described herein, a particular advantage of a filamentous phage in the present invention is that the DNA molecule present in the phage particle and encoding one or both of the members of the heterodimeric receptor can be segregated from other DNA molecules present in the library on the basis of the presence of the particular expressed fusion protein the surface of the phage particle.

Isolation (segregation) of a DNA molecule encoding one or both members of a heterodimeric receptor is conducted by segregation of the filamentous phage particle containing the gene or genes of interest away from the population of other phage particles comprising the library. Segregation of phage particles involves the physical separation and propagation of individual phage particles away from other particles in the library. Methods for physical separation of filamentous phage particles to produce individual particles, and the propagation of the individual particles to form populations of progeny phage derived from the individual segregated particle are well known in the filamentous phage arts.

A preferred separation method involves the identification of the expressed heterodimer on the surface of the phage particle by means of a ligand binding specificity between the phage particle and a preselected ligand. Exemplary and preferred is the use of "panning" methods whereby a suspension of phage particles is contacted with a solid phase ligand (antigen) and allowed to specifically bind (or immunoreact where the heterodimer includes an immunoglobulin variable domain). After binding, non-bound particles are washed off the solid phase, and the bound phage particles are those that contain ligand-specific heterodimeric receptor (heterodimer) on their surface. The bound particles can then be recovered by elution of the bound particle from the solid phase, typically by the use of aqueous solvents having high ionic strength sufficient to disrupt the receptor-ligand binding interaction.

An alternate method for separating a phage particle based on the ligand specificity of the surface-expressed heterodimer from a population of particles is to precipitate the phage particles from the solution phase by crosslinkage with the ligand. An exemplary and preferred crosslinking and precipitation method is described in detail in Example 4c.

The use of the above particle segregation methods provides a means for screening a population of filamentous phage particles present in a phage library of this invention. As applied to a phage library, screening can be utilized to enrich the library for one or more particles that express a heterodimer having a preselected ligand binding specificity. Where the library is designed to contain multiple species of heterodimers that all have some detectable measure of ligand binding activity, but differ in protein structure, antigenicity, ligand binding affinity or avidity, and the like, the screening methods can be utilized sequentially to first produce a library enriched for a preselected binding specificity, and then to produce a second library further

enriched by further screening comprising one or more isolated phage particles. Methods for measuring ligand binding activities, antigenicity and the like interactions between a ligand and a receptor are generally well known and are not discussed further as they are not essential features of the present invention.

Thus, in one embodiment, a phage library is a population of particles enriched for a preselected ligand binding specificity.

In another embodiment, a phage library comprises a population of particles wherein each particle contains at least one fusion protein of this invention on the surface of the phage particle. The actual amount of fusion protein present on the surface of a phage particle depends, in part, on the choice of coat protein membrane anchor present in the fusion protein. Where the anchor is derived from cpIII, there are typically about 1 to 4 fusion proteins per phage particle. Where the anchor is derived from the more preferred cpVIII, there is the potential for hundreds of fusion proteins on the particle surface depending on the growth conditions and other factors as discussed herein. Preferably, a phage particle in a library contains from about 10 to about 500 cpVIII-derived fusion proteins on the surface of each particle, and more preferably about 20 to 50 fusion proteins per particle. Exemplary amounts of surface fusion protein are shown by the electron micrographs described in Example 4a that describe particles having about 20 to 24 cpVIII-derived fusion proteins per particle.

In another embodiment, the present invention contemplates a population of phage particles that are the progeny of a single particle, and therefor all express the same heterodimer on the particle surface. Such a population of phage are homogeneous and clonally derived, and therefore provide a source for expressing large quantities of a particular fusion protein. An exemplary clonally homogeneous phage population is described in Example 4.

A filamentous phage particle in a library of this invention is produced by standard filamentous phage particle preparation methods and depends on the presence in a DNA expression vector of this invention of a filamentous phage origin of replication as described herein to provide the signals necessary for (1) production of a single-stranded filamentous phage replicative form and (2) packaging of the replicative form into a filamentous phage particle. Such a DNA molecule can be packaged when present in a bacterial cell host upon introduction of genetic complementation to provide the filamentous phage proteins required for production of infectious phage particles. A typical and preferred method for genetic complementation is to infect a bacterial host cell containing a DNA expression vector of this invention with a helper filamentous phage, thereby providing the genetic elements required for phage particle assembly. Exemplary helper rescue methods are described herein at Example 2, and described by Short et al., *Nuc. Acids Res.*, 16:7583–7600 (1988).

The level of heterodimeric receptor captured on the surface of a filamentous phage particle during the process of phage particle extrusion from the host cell can be controlled by a variety of means. In one embodiment, the levels of fusion proteins are controlled by the use of strong promoters in the first and second cistrons for expressing the polypeptides, such that transcription of the fusion protein cistrons occurs at a relative rate greater than the rate of transcription of the cpVIII gene on the helper phage. In another embodiment, the helper phage can have an amber mutation in the gene for expressing cpVIII, such that less wild-type cpVIII is transcribed in the host cell than fusion proteins, thereby leading to increased ratios of fusion protein compared to cpVIII during the extrusion process.

In another embodiment, the amount of heterodimeric receptor on the phage particle surface can be controlled by controlling the timing between expression of fusion proteins and the superinfection by helper phage. After introduction of the expression vector, longer delay times before the addition of helper phage will allow for increased accumulation of the fusion proteins in the host cell.

EXAMPLES

The following examples are intended to illustrate, but not limit, the scope of the invention.

1. Construction of a Dicistronic Expression Vector for Producing a Heterodimeric Receptor on Phage Particles

To obtain a vector system for generating a large number of Fab antibody fragments that can be screened directly, expression libraries in bacteriophage Lambda have previously been constructed as described in Huse et al., *Science*, 246:1275–1281 (1989). These systems did not contain design features that provide for the expressed Fab to be targeted to the surface of a filamentous phage particle.

The main criterion used in choosing a vector system was the necessity of generating the largest number of Fab fragments which could be screened directly. Bacteriophage Lambda was selected as the starting point to develop an expression vector for three reasons. First, in vitro packaging of phage DNA was the most efficient method of reintroducing DNA into host cells. Second, it was possible to detect protein expression at the level of single phage plaques. Finally, the screening of phage libraries typically involved less difficulty with nonspecific binding. The alternative, plasmid cloning vectors, are only advantageous in the analysis of clones after they have been identified. This advantage was not lost in the present system because of the use of a dicistronic expression vector such as pCombVIII, thereby permitting a plasmid containing the heavy chain, light chain, or Fab expressing inserts to be excised.

a. Construction of Dicistronic Expression Vector pCOMB

(i) Preparation of Lambda Zap™ II

Lambda Zap™ II is a derivative of the original Lambda Zap (ATCC #40.298) that maintains all of the characteristics of the original Lambda Zap including 6 unique cloning sites, fusion protein expression, and the ability to rapidly excise the insert in the form of a phagemid (Bluescript SK–), but lacks the SAM 100 mutation, allowing growth on many Non-Sup F strains, including XL1-Blue. The Lambda Zap™ II was constructed as described in Short et al., *Nuc. Acids Res.*, 16:7583–7600, 1988, by replacing the Lambda S gene contained in a 4254 base pair (bp) DNA fragment produced by digesting Lambda Zap with the restriction enzyme Nco I. This 4254 bp DNA fragment was replaced with the 4254 bp DNA fragment containing the Lambda S gene isolated from Lambda gt10 (ATCC #40.179) after digesting the vector with the restriction enzyme Nco I. The 4254 bp DNA fragment isolated from lambda gt10 was ligated into the original Lambda Zap vector using T4 DNA ligase and standard protocols such as those described in *Current Protocols in Molecular Biology*, Ausubel et al., eds., John Wiley and Sons, NY, 1987, to form Lambda Zap™ II.

(ii) Preparation of Lambda Hc2

To express a plurality of V_H-coding DNA homologs in an *E. coli* host cell, a vector designated Lambda Hc2 was constructed. The vector provided the following: the capacity to place the V_H-coding DNA homologs in the proper reading frame; a ribosome binding site as described by Shine et al., *Nature*, 254:34, 1975; a leader sequence directing the expressed protein to the periplasmic space designated the pelB secretion signal; a polynucleotide sequence that coded for a known epitope (epitope tag); and also a polynucleotide that coded for a spacer protein between the V_H-coding DNA homolog and the polynucleotide coding for the epitope tag. Lambda Hc2 has been previously described by Huse et al., *Science*, 246:1275–1281 (1989).

TABLE 3

SEQ. ID. NO.		
(22)	N1)	5' GGCCGCAAATTCTATTTCAAGGAGACAGTCAT 3'
(23)	N2)	5' AATGAAATACCTATTGCCTACGGCAGCCGCTGGATT 3'
(24)	N3)	5' GTTATTACTCGCTGCCCAACCAGCCATGGCCC 3'
(25)	N6)	5' CAGTTTCACCTGGGCCATGGCTGGTTGGG 3'
(26)	N7)	5' CAGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAG 3'
(27)	N8)	5' GTATTTCATTATGACTGTCTCCTTGAAATAGAATTTGC 3'
(28)	N9-4)	5' AGGTGAAACTGCTCGAGATTTCTAGACTAGTTACCCGTAC 3'
(29)	N10-5)	5' CGGAACGTCGTACGGGTAACTAGTCTAGAAATCTCGAG 3'
(30)	N11)	5' GACGTTCCGGACTACGGTTCTTAATAGAATTCG 3'
(31)	N12)	5' TCGACGAATTCTATTAAGAACCGTAGTC 3'

To prepare Lambda Hc2, a synthetic DNA sequence containing all of the above features was constructed by designing single stranded polynucleotide segments of 20–40 bases that would hybridize to each other and form the double stranded synthetic DNA sequence shown in FIG. 3. The individual single-stranded polynucleotide segments are shown in Table 3.

Polynucleotides N2, N3, N9-4, N11, N10-5, N6, N7 and N8 (Table 3) were kinased by adding 1 µl of each polynucleotide 0.1 micrograms/microliter (µg/µl) and 20 units of T_4 polynucleotide kinase to a solution containing 70 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 10 mM beta-mercaptoethanol, 500 micrograms per milliliter (µg/ml) bovine serum albumin (BSA). The solution was maintained at 37 degrees Centigrade (37° C.) for 30 minutes and the reaction stopped by maintaining the solution at 65° C. for 10 minutes. The two end polynucleotides, 20 ng of polynucleotides N1 and polynucleotides N12, were added to the above kinasing reaction solution together with 1/10 volume of a solution containing 20.0 mM Tris-HCl, pH 7.4, 2.0 mM $MgCl_2$ and 50.0 mM NaCl. This solution was heated to 70° C. for 5 minutes and allowed to cool to room temperature, approximately 25° C., over 1.5 hours in a 500 ml beaker of water. During this time period all 10 polynucleotides annealed to form the double stranded synthetic DNA insert shown in FIG. 3. The individual polynucleotides were covalently linked to each other to stabilize the synthetic DNA insert by adding 40 µl of the above reaction to a solution containing 50 mM Tris-HCl, pH 7.5, 7 mM $MgCl_2$, 1 mM DTT, 1 mM adenosine triphosphate (ATP) and 10 units of T4 DNA ligase. This solution was maintained at 37° C. for 30 minutes and then the T4 DNA ligase was inactivated by maintaining the solution at 65° C. for 10 minutes. The end polynucleotides were kinased by mixing 52 µl of the above reaction, 4 µl of a solution containing 10 mM ATP and 5 units of T4 polynucleotide kinase. This solution was maintained at 37° C. for 30 minutes and then the T4 polynucleotide kinase was inactivated by maintaining the solution at 65° C. for 10 minutes.

The completed synthetic DNA insert was ligated directly into the Lambda Zap™ II vector described in Example 1a(i) that had been previously digested with the restriction enzymes, Not I and Xho I. The ligation mixture was packaged according to the manufacture's instructions using Gigapack II Gold packing extract available from Stratagene, La Jolla, Calif. The packaged ligation mixture was plated on XL1-Blue cells (Stratagene). Individual lambda plaques were cored and the inserts excised according to the in vivo excision protocol for Lambda Zap™ II provided by the manufacturer (Stratagene). This in vivo excision protocol moved the cloned insert from the Lambda Hc2 vector into a phagemid vector to allow easy for manipulation and sequencing. The accuracy of the above cloning steps was confirmed by sequencing the insert using the Sanger dideoxy method described in by Sanger et al., *Proc. Natl. Acad. Sci. USA*, 74:5463–5467, (1977) and using the manufacture's instructions in the AMV Reverse Transcriptase ^{35}S-ATP sequencing kit (Stratagene). The sequence of the resulting double-stranded synthetic DNA insert in the V_H expression vector (Lambda Hc2) is shown in FIG. 3. The sequence of each strand (top and bottom) of Lambda Hc2 is listed in the sequence listing as SEQ. ID. NO. 1 and SEQ. ID. NO. 2, respectively. The resultant Lambda Hc2 expression vector is shown in FIG. 4.

(iii) Preparation of Lambda Lc2

To express a plurality of V_L-coding DNA homologs in an *E. coli* host cell, a vector designated Lambda Lc2 was constructed having the capacity to place the V_L-coding DNA homologs in the proper reading frame, provided a ribosome binding site as described by Shine et al., *Nature*, 254:34 (1975), provided the pelB gene leader sequence secretion signal that has been previously used to successfully secrete Fab fragments in *E. coli* by Lei et al., *J. Bac.*, 169:4379 (1987) and Better et al., *Science*, 240:1041 (1988), and also provided a polynucleotide containing a restriction endonuclease site for cloning. Lambda Lc2 has been previously described by Huse et al., *Science*, 246:1275–1281 (1989).

A synthetic DNA sequence containing all of the above features was constructed by designing single stranded polynucleotide segments of 20–60 bases that would hybridize to each other and form the double stranded synthetic DNA sequence shown in FIG. 5. The sequence of each individual single-stranded polynucleotide segment (01–08) within the double stranded synthetic DNA sequence is shown in Table 4.

Polynucleotides 02, 03, 04, 05, 06 and 07 (Table 4) were kinased by adding 1 µl (0.1 ug/µl) of each polynucleotide

and 20 units of T_4 polynucleotide kinase to a solution containing 70 mM Tris-HCl. pH 7.6. 10 mM MgCl. 5 mM DTT. 10 mM beta-mercaptoethanol. 500 mg/ml of BSA. The solution was maintained at 37° C. for 30 minutes and the reaction stopped by maintaining the solution at 65° C. for 10 minutes. The 20 ng each of the two end polynucleotides. 01 and 08. were added to the above kinasing reaction solution together with 1/10 volume of a solution containing 20.0 mM Tris-HCl. pH 7.4. 2.0 mM MgCl and 15.0 mM sodium chloride (NaCl). This solution was heated to 70° C. for 5 minutes and allowed to cool to room temperature. approximately 25° C.. over 1.5 hours in a 500 ml beaker of water. During this time period all 8 polynucleotides annealed to form the double stranded synthetic DNA insert shown in FIG. 5. The individual polynucleotides were covalently linked to each other to stabilize the synthetic DNA insert by adding 40 µl of the above reaction to a solution containing 50 ml Tris-HCl. pH 7.5. 7 ml MgCl. 1 mm DTT. 1 mm ATP and 10 units of T4 DNA ligase. This solution was maintained at 37° C. for 30 minutes and then the T4 DNA ligase was inactivated by maintaining the solution at 65° C. for 10 minutes. The end polynucleotides were kinased by mixing 52 µl of the above reaction. 4 µl of a solution containing 10 mM ATP and 5 units of T4 polynucleotide kinase. This solution was maintained at 37° C. for 30 minutes and then the T4 polynucleotide kinase was inactivated by maintaining the solution at 65° C. for 10 minutes.

TABLE 4

SEQ. ID. NO.		
(32)	01)	5' TGAATTCTAAACTAGTCGCCAAGGAGACAGTCAT 3'
(33)	02)	5' AATGAAATACCTATTGCCTACGGCAGCCGCTGGATT 3'
(34)	03)	5' GTTATTACTCGCTGCCCAACCAGCCATGGCC 3'
(35)	04)	5' GAGCTCGTCAGTTCTAGAGTTAAGCGGCCG 3'
(36)	05)	5' GTATTTCATTATGACTGTCTCCTTGGCGACTAGTTTAGAA-TTCAAGCT 3'
(37)	06)	5' CAGCGAGTAATAACAATCCAGCGGCTGCCGTAGGCAATAG 3'
(38)	07)	5' TGACGAGCTCGGCCATGGCTGGTTGGG 3'
(39)	08)	5' TCGACGGCCGCTTAACTCTAGAAC 3'

The completed synthetic DNA insert was ligated directly into the Lambda Zap™ II vector described in Example 1(a)(i) that had been previously digested with the restriction enzymes Sac I and Xho I. The ligation mixture was packaged according to the manufacture's instructions using Gigapack II Gold packing extract (Stratagene). The packaged ligation mixture was plated on XL1-Blue cells (Stratagene). Individual lambda plaques were cored and the inserts excised according to the in vivo excision protocol for Lambda Zap™ II provided by the manufacturer (Stratagene). This in vivo excision protocol moved the cloned insert from the Lambda Lc2 vector into a plasmid phagemid vector allow for easy manipulation and sequencing. The accuracy of the above cloning steps was confirmed by sequencing the insert using the manufacture's instructions in the AMV Reverse Transcriptase ^{35}S-dATP sequencing kit (Stratagene). The sequence of the resulting Lc2 expression vector (Lambda Lc2) is shown in FIG. 5. Each strand is separately listed in the Sequence Listing as SEQ. ID. NO. 3 and SEQ. ID. NO. 4. The resultant Lc2 vector is schematically diagrammed in FIG. 6.

A preferred vector for use in this invention. designated Lambda Lc3. is a derivative of Lambda Lc2 prepared above. Lambda Lc2 contains a Spe I restriction site (ACTAGT) located 3' to the EcoR I restriction site and 5' to the Shine-Dalgarno ribosome binding site as shown in the sequence in FIG. 5 and in SEQ. ID. NO. 3. A Spe I restriction site is also present in Lambda Hc2 as shown in FIGS. 3 and 4 and in SEQ. ID. NO. 1. A combinatorial vector. designated pComb. was constructed by combining portions of Lambda Hc2 and Lc2 together as described in Example 1a(iv) below. The resultant combinatorial pComb vector contained two Spe I restriction sites. one provided by Lambda Hc2 and one provided by Lambda Lc2. with an EcoR I site in between. Despite the presence of two Spe I restriction sites. DNA homologs having Spe I and EcoR I cohesive termini were successfully directionally ligated into a pComb expression vector previously digested with Spe I and EcoR I as described in Example 1b below. The proximity of the EcoR I restriction site to the 3' Spe I site. provided by the Lc2 vector. inhibited the complete digestion of the 3' Spe I site. Thus. digesting pComb with Spe I and EcoR I did not result in removal of the EcoR I site between the two Spe I sites.

The presence of a second Spe I restriction site may be undesirable for ligations into a pComb vector digested only with Spe I as the region between the two sites would be eliminated. Therefore. a derivative of Lambda Lc2 lacking the second or 3' Spe I site. designated Lambda Lc3. is produced by first digesting Lambda Lc2 with Spe I to form a linearized vector. The ends are filled in to form blunt ends which are ligated together to result in Lambda Lc3 lacking a Spe I site. Lambda Lc3 is a preferred vector for use in constructing a combinatorial vector as described below.

(iv) Preparation of pComb

Phagemids were excised from the expression vectors Lambda Hc2 or Lambda Lc2 using an in vivo excision protocol described above. Double stranded DNA was prepared from the phagemid-containing cells according to the methods described by Holmes et al., *Anal. Biochem.*, 114:193 (1981). The phagemids resulting from in vivo excision contained the same nucleotide sequences for antibody fragment cloning and expression as did the parent vectors. and are designated phagemid Hc2 and Lc2. corresponding to Lambda Hc2 and Lc2. respectively.

For the construction of combinatorial phagemid vector pComb. produced by combining portions of phagemid Hc2 and phagemid Lc2. phagemid Hc2 was first digested with Sac I to remove the restriction site located 5' to the LacZ promoter. The linearized phagemid was then blunt ended with T4 polymerase and ligated to result in a Hc2 phagemid lacking a Sac I site. The modified Hc2 phagemid and the Lc2 phagemid were then separately restriction digested with Sca I and EcoR I and the linearized phagemids were ligated together at their respective cohesive ends. The ligated phagemid vector was then inserted into an appropriate bacterial host and transformants were selected on the antibiotic ampicillin.

Selected ampicillin resistant transformants were screened for the presence of two Not I sites. The resulting ampicillin resistant combinatorial phagemid vector was designated pComb, the schematic organization of which is shown in FIG. 7. The resultant combinatorial vector, pComb, consisted of a DNA molecule having two cassettes to express two fusion proteins and having nucleotide residue sequences for the following operatively linked elements listed in a 5' to 3' direction: a first cassette consisting of an inducible LacZ promoter upstream from the LacZ gene; a Not I restriction site; a ribosome binding site; a pelB leader; a spacer; a cloning region bordered by a 5' Xho and 3' Spe I restriction site; a decapeptide tag followed by expression control stop sequences; an EcoR I restriction site located 5' to a second cassette consisting of an expression control ribosome binding site; a pelB leader; a spacer region; a cloning region bordered by a 5' Sac I and a 3' Xba I restriction site followed by expression control stop sequences and a second Not I restriction site.

A preferred combinatorial vector for use in this invention, designated pComb2, is constructed by combining portions of phagemid Hc2 and phagemid Lc3 as described above for preparing pComb. The resultant combinatorial vector, pComb2, consists of a DNA molecule having two cassettes identical to pComb to express two fusion proteins identically to pComb except that a second Spe I restriction site in the second cassette is eliminated.

b. Construction of Vectors pCombVIII and pCombIII for Expressing Fusion Proteins Having a Bacteriophage Coat Protein Membrane Anchor

Because of the multiple endonuclease restriction cloning sites, the pComb phagemid expression vector prepared above is a useful cloning vehicle for modification for the preparation of an expression vector of this invention. To that end, pComb is digested with EcoR I and Spe I followed by phosphatase treatment to produce linearized pComb.

(i) Preparation of pCombVIII

A PCR product produced in Example 2g and having a nucleotide sequence that defines a filamentous bacteriophage coat protein VIII (cpVIII) membrane anchor domain and cohesive Spe I and EcoR I termini was admixed with the linearized pComb to form a ligation admixture. The cpVIII-membrane anchor-encoding PCR fragment was directionally ligated into the pComb phagemid expression vector at corresponding cohesive termini, that resulted in forming pCombVIII (also designated pComb8). pCombVIII contains a cassette defined by the nucleotide sequence shown in SEQ. ID. NO. 116 from nucleotide base 1 to base 208, and contains a pelB secretion signal operatively linked to the cpVIII membrane anchor.

A preferred phagemid expression vector for use in this invention, designated either pComb2-VIII or pComb2-8, is prepared as described above by directionally ligating the cpVIII membrane anchor-encoding PCR fragment into a pComb2 phagemid expression vector via Spe I and EcoR I cohesive terminii. The pComb2-8 has only one Spe I restriction site.

(ii) Preparation of pCombIII

A separate phagemid expression vector was constructed using sequences encoding bacteriophage cpIII membrane anchor domain. A PCR product defining the cpIII membrane anchor and Spe I and EcoR I cohesive termini was prepared as described for cpVIII, the details of which are described in Example 2g. The cpIII-derived PCR product was then ligated into linearized pComb vector to form the vector pCombIII (also designated pComb3).

A preferred phagemid expression vector for use in this invention, designated either pComb2-III or pComb2-3, is prepared as described above by directionally ligating the cpIII membrane anchor-encoding PCR fragment into a pComb2 phagemid expression vector via Spe I and Spe I cohesive terminii. The pComb2-3 has only one Spe I restriction site.

c. Construction of pCBAK Vectors Having a Chloramphenicol Resistance Marker

In order to utilize a different selectable marker gene, such as chloramphenicol acetyl transferase (CAT), for the selection of bacteria transformed with a vector of this invention, expression vectors based on pComb were developed having a gene encoding CAT and are designated pCBAK vectors. The pCBAK vectors are prepared by combining portions of pCB and pComb.

(i) Preparation of pCB

pBlueScript phagemid vectors, pBC SK(−) and pBS SK(−). (Stratagene), were modified and combined to generate a third vector designated pCB as described below.

pBC SK(−), which contains a chloramphenicol resistance selectable marker gene, was digested with Bst BI and blunt ended with T4 polymerase. A second digestion with Pvu I allowed for the removal of a 1 kilobase (kb) fragment leaving a 2.4 kb linearized vector which retained the CAT selectable resistance marker gene, an inducible LacZ promoter upstream from the LacZ gene and a ColE1 origin region. The 2.4 kb fragment was recovered. The pBS SK(−) vector was digested with Aat II and blunt ended with T4 polymerase. A second digestion with Pvu I allowed for the isolation of an 800 base pair (bp) fragment containing the f1 origin of replication. Ligation of the pBS derived 800 bp f1 fragment with the 2.4 kb pBC fragment created a pCB precursor vector containing a Sac I site, an f1 origin of replication, a CAT selectable resistance marker gene, ColE1 origin, a multiple cloning site (MCS) flanked by T_3 and T_7 promoters, and an inducible LacZ promoter upstream from LacZ gene.

The pCB precursor vector was then digested with Sac I and blunt-ended with T4 polymerase. The T4 polymerase-treated pCB vector was then religated to form pCB vector and is lacking a Sac I site.

(ii) Preparation of pCBAK0

The pCB vector containing the CAT selectable resistance marker gene was digested with Sac II and Apa I and treated with phosphatase to prevent religation and to form linearized pCB vector. The pComb vector prepared in Example 1(a)(iv) was restriction digested with Sac II and Apa I to release a fragment containing nucleotide residue sequences starting 5' to the LacZ promoter and extending past the 3' end of the second Not I site. The Sac II and Apa I pComb DNA fragment was then directionally ligated into the similarly digested pCB vector to form phagemid expression vector pCBAK0. Preferred pCBAK expression vectors are constructed with pComb2. The resultant pCBAK expression vector contains only one Spe I restriction site.

(iii) Preparation of pCBAK8

To prepare a pCBAK-based phagemid expression vector which encodes a bacteriophage coat protein membrane anchor domain in the expressed fusion protein, pCB phagemid cloning vector prepared in Example 1c(ii) was linearized by digestion with Sac II and Apa I. The pCombVIII phagemid expression vector, prepared in Example 1b(i), was restriction digested with Sac II and Apa I to form a fragment containing a nucleotide residue sequence starting 5' to the LacZ promoter and extending past the 3' end of the

second Not 1 site. The fragment was directionally ligated into the linearized pCB cloning vector to form phagemid expression vector pCBAK8.

(iv) Preparation of pCBAK3

The phagemid expression vector, pCBAK3, for the expression of fusion protein having cpIII membrane anchor domains, was similarly constructed by directionally ligating the Sac II and Apa I restriction digested fragment from pCombIII with Sac II and Apa I linearized pCB cloning vector.

2. Construction of Dicistronic Expression Vectors for Expressing Anti-NPN Heterodimer on Phage Surfaces

In practicing this invention, the heavy (Fd consisting of V_H and C_H1) and light (kappa) chains (V_L, C_L) of antibodies are first targeted to the periplasm of *E. coli* for the assembly of heterodimeric Fab molecules. In order to obtain expression of antibody Fab libraries on a phage surface, the nucleotide residue sequences encoding either the Fd or light chains must be operatively linked to the nucleotide residue sequence encoding a filamentous bacteriophage coat protein membrane anchor. Two preferred coat proteins for use in this invention in providing a membrane anchor are VIII and III (cpVIII and cpIII, respectively). In the Examples described herein, methods for operatively linking a nucleotide residue sequence encoding a Fd chain to either cpVIII or cpIII membrane anchors in a fusion protein of this invention are described.

In a phagemid vector, a first and second cistron consisting of translatable DNA sequences are operatively linked to form a dicistronic DNA molecule. Each cistron in the dicistronic DNA molecule is linked to DNA expression control sequences for the coordinate expression of a fusion protein, Fd-cpVIII or Fd-cpIII, and a kappa light chain.

The first cistron encodes a periplasmic secretion signal (pelB leader) operatively linked to the fusion protein, either Fd-cpVIII or Fd-cpIII. The second cistron encodes a second pelB leader operatively linked to a kappa light chain. The presence of the pelB leader facilitates the coordinated but separate secretion of both the fusion protein and light chain from the bacterial cytoplasm into the periplasmic space.

The process described above is schematically diagrammed in FIG. 8. Briefly, the phagemid expression vector carries a chloramphenicol acetyl transferase (CAT) selectable resistance marker gene in addition to the Fd-cpVIII fusion and the kappa chain. The f1 phage origin of replication facilitates the generation of single stranded phagemid. The isopropyl thiogalactopyranoside (IPTG) induced expression of a dicistronic message encoding the Fd-cpVIII fusion (V_H, C_{H1}, cpVIII) and the light chain (V_L, C_L) leads to the formation of heavy and light chains. Each chain is delivered to the periplasmic space by the pelB leader sequence, which is subsequently cleaved. The heavy chain is anchored in the membrane by the cpVIII membrane anchor domain while the light chain is secreted into the periplasm. The heavy chain in the presence of light chain assembles to form Fab molecules. This same result can be achieved if, in the alternative, the light chain is anchored in the membrane via a light chain fusion protein having a membrane anchor and heavy chain is secreted via a pelB leader into the periplasm.

With subsequent infection of *E. coli* with a helper phage, as the assembly of the filamentous bacteriophage progresses, the coat protein VIII is incorporated along the entire length of the filamentous phage particles as shown in FIGS. 8 and 9. If cpIII is used, the accumulation occurs on the tail of the bacteriophage. The advantage of the utilization of membrane anchors from cpVIII over cpIII is two fold. Firstly, a multiplicity of binding sites, consisting of approximately 2700 cpVIII monomers assembled in a tubular array, exist along the particle surface. Secondly, the construct does not interfere with phage infectivity.

a. Polynucleotide Selection

The nucleotide sequences encoding the immunoglobulin protein CDR's are highly variable. However, there are several regions of conserved sequences that flank the V region domains of either the light or heavy chain, for instance, and that contain substantially conserved nucleotide sequences, i.e., sequences that will hybridize to the same primer sequence. Therefore, polynucleotide synthesis (amplification) primers that hybridize to the conserved sequences and incorporate restriction sites into the DNA homolog produced that are suitable for operatively linking the synthesized DNA fragments to a vector were constructed. More specifically, the primers are designed so that the resulting DNA homologs produced can be inserted into an expression vector of this invention in reading frame with the upstream translatable DNA sequence at the region of the vector containing the directional ligation means.

(i) V_H Primers

For amplification of the V_H domains, primers are designed to introduce cohesive termini compatible with directional ligation into the unique Xho I and Spe I sites of the phagemid Hc2 expression vector. For example, the 3' primer (primer 12A in Table 5), was designed to be complementary to the mRNA in the J_H region. In all cases, the 5' primers (primers 1–10, Table 5) were chosen to be complementary to the first strand cDNA in the conserved N-terminus region (antisense strand). Initially amplification was performed with a mixture of 32 primers (primer 1, Table 5) that were degenerate at five positions. Hybridoma mRNA could be amplified with mixed primers, but initial attempts to amplify mRNA from spleen yielded variable results. Therefore, several alternatives to amplification using the mixed 5' primers were compared.

The first alternative was to construct multiple unique primers, eight of which are shown in Table 5, corresponding to individual members of the mixed primer pool. The individual primers 2–9 of Table 5 were constructed by incorporating either of the two possible nucleotides at three of the five degenerate positions.

The second alternative was to construct a primer containing inosine (primer 10, Table 5) at four of the variable positions based on the published work of Takahashi, et al., *Proc. Natl. Acad. Sci. (U.S.A.)*, 82:1931–1935, (1985) and Ohtsuka et al., *J. Biol. Chem.*, 260:2605–2608, (1985). This primer has the advantage that it is not degenerate and, at the same time minimizes the negative effects of mismatches at the unconserved positions as discussed by Martin et al., *Nuc. Acids Res.*, 13:8927 (1985). However, it was not known if the presence of inosine nucleotides would result in incorporation of unwanted sequences in the cloned V_H regions. Therefore, inosine was not included at the one position that remains in the amplified fragments after the cleavage of the restriction sites. As a result, inosine was not in the cloned insert.

Additional V_H amplification primers including the unique 3' primer were designed to be complementary to a portion of the first constant region domain of the gamma 1 heavy chain mRNA (primers 16 and 17, Table 5). These primers will produce DNA homologs containing polynucleotides coding for amino acids from the V_H and the first constant region

domains of the heavy chain. These DNA homologs can therefore be used to produce Fab fragments rather than an F_V.

Additional unique 3' primers designed to hybridize to similar regions of another class of immunoglobulin heavy chain such as IgM. IgE and IgA are contemplated. Other 3' primers that hybridize to a specific region of a specific class of CH_1 constant region and are adapted for transferring the V_H domains amplified using this primer to an expression vector capable of expressing those V_H domains with a different class of heavy or light chain constant region are also contemplated.

As a control for amplification from spleen or hybridoma mRNA. a set of primers hybridizing to a highly conserved region within the constant region IgG. heavy chain gene were constructed. The 5' primer (primer 11. Table 5) is complementary to the cDNA in the C_H2 region whereas the 3' primer (primer 13. Table 5) is complementary to the mRNA in the C_H3 region. It is believed that no mismatches were present between these primers and their templates.

The primers used for amplification of heavy chain Fd fragments for construction of Fabs are shown at least in Table 5. Amplification was performed in eight separate reactions. each containing one of the 5' primers (primers 2–9) and one of the 3' primers (primer 16). The remaining 5' primers that have been used for amplification in a single reaction are either a degenerate primer (primer 1) or a primer that incorporates inosine at four degenerate positions (primer 10. Table 5. and primers 17 and 18. Table 6). The remaining 3' primer (primer 14. Table 6) has been used to construct F_V fragments. Many of the 5' primers incorporate a Xho I site. and the 3' primers incorporate a Spe I restriction site for insertion of the V_H DNA homolog into the phagemid Hc2 expression vector (FIG. 4).

V_H amplification primers designed to amplify human heavy chain variable regions are shown in Table 6. One of the 5' heavy chain primer contains inosine residues at degenerate nucleotide positions allowing a single primer to hybridize to a large number of variable region sequences. Primers designed to hybridize to the constant region sequences of various IgG mRNAs are also shown in Table 6.

(ii) V_L Primers

The nucleotide sequences encoding the V_L CDRs are highly variable. However. there are several regions of conserved sequences that flank the V_L CDR domains including the J_L. V_L framework regions and V_L leader/promotor. Therefore. amplification primers were constructed that hybridized to the conserved sequences and incorporate restriction sites that allow cloning the amplified fragments into the phagemid Lc2 vector cut with Sac I and Xba I.

For amplification of the V_L CDR domains. the 5' primers (primers 1–8 in Table 6) were designed to be complementary to the first strand cDNA in the conserved N-terminus region. These primers also introduced a Sac I restriction endonuclease site to allow the V_L DNA homolog to be cloned into the phagemid Lc2 expression vector. The 3' V_L amplification primer (primer 9 in Table 6) was designed to be complementary to the mRNA in the J_L regions and to introduce the Xba I restriction endonuclease site required to insert the V_L DNA homolog into the phagemid Lc2 expression vector (FIG. 6).

Additional 3' V_L amplification primers were designed to hybridize to the constant region of either kappa or lambda mRNA (primers 10 and 11 in Table 6). These primers allow a DNA homolog to be produced containing polynucleotide sequences coding for constant region amino acids of either kappa or lambda chain. These primers make it possible to produce an Fab fragment rather than an F_V.

The primers used for amplification of kappa light chain sequences for construction of Fabs are shown at least in Table 6. Amplification with these primers was performed in 5 separate reactions. each containing one of the 5' primers (primers 3–6. and 12) and one of the 3' primers (primer 13). The remaining 3' primer (primer 9) has been used to construct F_V fragments. The 5' primers contain a Sac I restriction site and the 3' primers contain a Xba I restriction site.

V_L amplification primers designed to amplify human light chain variable regions of both the lambda and kappa isotypes are also shown in Table 6.

All primers and synthetic polynucleotides described herein. including those shown in Tables 3–7 were either purchased from Research Genetics in Huntsville. Ala. or synthesized on an Applied Biosystems DNA synthesizer. model 381A. using the manufacturer's instruction.

TABLE 5

(1)	5'AGGT(C/G)(C/A)A(G/A)CT(G/T)CTCGAGTC(T/A)GG 3'	degenerate 5' primer for the amplification of mouse and human heavy chain variable regions (V_H)
(2)	5'AGGTCCAGCTGCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(3)	5'AGGTCCAGCTGCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(4)	5'AGGTCCAGCTTCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(5)	5'AGGTCCAGCTTCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(6)	5'AGGTCCAACTGCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(7)	5'AGGTCCAACTGCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(8)	5'AGGTCCAACTTCTCGAGTCTGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(9)	5'AGGTCCAACTTCTCGAGTCAGG 3'	Unique 5' primer for the amplification of mouse and human V_H
(10)	5'AGGTNNANCTNCTCGAGTC(T/A)GG 3'	5' degenerate primer containing inosine at 4 degenerate positions for amplification of mouse V_H
(11)	5'GCCCAAGGATGTGCTCACC 3'	5' primer for amplification in the C_H2 region of mouse IgG1

TABLE 5-continued

(12)	5'CTATTAGAATTCAACGGTAACAGTGGTGCCTTGGCCCCA 3'	3' primer for amplification of V_H and introducing a 3' Eco RI site
(12A)	5'CTATTAACTAGTAACGGTAACAGTGGTGCCTTG CCCCA 3'	3' primer for amplification of V_H using 3' Spe I site
(13)	5'CTCAGTATGGTGGTTGTGC 3'	3' primer for amplification in the C_H3 region of mouse IgG1
(14)	5'GCTACTAGTTTTGATTTCCACCTTGG 3'	3' primer for amplification of mouse kappa light chain variable regions (V_L)
(15)	5'CAGCCATGGCCGACATCCAGATG 3'	5' primer for amplification of mouse kappa light chain variable regions
(16)	5'AATTTTACTAGTCACCTTGGTGCTGCTGGC 3'	Unique 3' primer for amplification of V_H including part of the mouse gamma 1 first constant region
(17)	5'TATGCAACTAGTACAACCACAATCCCTGGGCACAATTTT 3'	Unique 3' primer for amplification of Fd including part of mouse IgG1 first constant region and hinge region
(18)	5'AGGCTTACTAGTACAATCCCTGGGCACAAT 3'	3' primer for amplifying mouse Fd including part of the mouse IgG first constant region and part of the hinge region

TABLE 6

(1)	5' CCAGTTCCGAGCTCGTTGTGACTCAGGAATCT 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(2)	5' CCAGTTCCGAGCTCGTGTTGACGCAGCCGCCC 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(3)	5' CCAGTTCCGAGCTCGTGCTCACCCAGTCTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(4)	5' CCAGTTCCGAGCTCCAGATGACCCAGTCTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(5)	5' CCAGATGTGAGCTCGTGATGACCCAGACTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(6)	5' CCAGATGTGAGCTCGTCATGACCCAGTCTCCA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(7)	5' CCAGATGTGAGCTCTTGATGACCCAAACTCAA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(8)	5' CCAGATGTGAGCTCGTGATAACCCAGGATGAA 3'	Unique 5' primer for the amplication of kappa light chain variable regions
(9)	5' GCAGCATTCTAGAGTTTCAGCTCCAGCTTGCC 3'	Unique 3' primer for amplification of kappa light chain variable regions
(10)	5' CCGCCGTCTAGAACACTCATTCCTGTTGAAGCT 3'	Unique 3' primer for mouse kappa light chain amplification including the constant region
(11)	5' CCGCCGTCTAGAACATTCTGCAGGAGACAGACT 3'	Unique 3' primer for mouse lambda light chain amplification including the constant region
(12)	5' CCAGTTCCGAGCTCGTGATGACACAGTCTCCA 3'	Unique 5' primer for V_L amplification
(13)	5' GCGCCGTCTAGAATTAACACTCATTCCTGTTGAA 3'	Unique 3' primer for amplification of kappa light chain
(14)	5' CTATTAACTAGTAACGGTAACAGTGGTGCCTTGCCCCA 3'	Unique 3' primer for amplification of mouse F_V
(15)	5' AGGCTTACTAGTACAATCCCTGGGCACAAT 3'	Unique 3' primer for amplification of mouse IgG Fd
(16)	5' GCCGCTCTAGAACACTCATTCCTGTTGAA 3'	Unique 3' primer for amplification of mouse kappa light chain
(17)	5' AGGTIIAICTICTCGAGTCTGC 3'	Degenerate 5' primer containing inosine at 4 degenerate positions for amplifying mouse V_H
(18)	5' AGGTIIAICTICTCGAGTCAGC 3'	Degenerate 5' primer containing inosine at 4 degenerate positions for amplifying mouse V_H
(19)	5' GTGCCAGATGTGAGCTCGTGATGACCCAGTCTCCA 3'	Unique 5' primer for human and mouse kappa V_L amplification
(20)	5' TCCTTCTAGATTACTAACACTCTCCCCTGTTGAA 3'	Unique 3' primer for kappa V_L amplification
(21)	5' GCATTCTAGACTATTATGAACATTCTGTAGGGGC 3'	Unique 3' primer for human, mouse and rabbit lambda V_L amplification
(22)	5' CTGCACAGGGTCCTGGGCCGAGCTCGTGGTGACTCAG 3'	Unique 5' primer for human lambda V_L amplification
(23)	5' AGITGCAIITGCTCGAGTCTGG 3'	5' degenerate primer for human V_H amplification containing inosine at 3 degenerate positions
(24)	5' GTGGGCATGTGTGAGTTGTGTCACTAGTTGGGGTTTTGAGCTC 3'	Unique 3' primer for human V_H amplification
(25)	5' AGCATCACTAGTACAAGATTTGGGCTC 3'	Unique 3' primer for human IgG1 Fd

TABLE 6-continued

		amplification
(26)	5' AGCATCACTAGTACAAGATTTGGGCTC 3'	Unique 3' primer for amplification of human variable regions (V_H)
(27)	5' AGGTGCAGCTGCTCGAGTCGGG 3'	Unique 3' primer for amplification of human variable regions (V_H)
(28)	5' AGGTGCAACTGCTCGAGTCTGG 3'	Unique 3' primer for amplification of human variable regions (V_H)
(29)	5' AGGTGCAACTGCTCGAGTCGGG 3'	Unique 3' primer for amplification of human variable regions (V_H)
(30)	5' TCCTTCTAGATTACTAACACTCTCCCCTGTTGAA 3'	3' primer in human kappa light chain constant region
(31)	5' CTGCACAGGGTCCTGGGCCGAGCTCGTGGTGACTCAG 3'	5' primer for amplification of human lambda light chain variable regions
(32)	5' GCATTCTAGACTATTAACATTCTGTAGGGGC 3'	3' primer in human lambda light chain constant region
(33)	5' ACCCAAGGACACCCTCATG 3'	Control primer hybridizing to the human CH_2 region
(34)	5' CTCAGTATGGTGGTTGTGC 3'	Control primer hybridizing to the human CH_3 region
(35)	5' GTCTCACTAGTCTCCACCAAGGGCCCATCGGTC 3'	5' primer for amplifying human IgG heavy chain first constant region
(36)	5' ATATACTAGTGAGACAGTGACCAGGGTTCCTTGGCCCCA 3'	3' primer for amplifying human heavy chain variable regions
(37)	5' ACGTCTAGATTCCACCTTGGTCCC 3'	3' primer for amplifying human kapp chain variable regions
(38)	5' GCATACTAGTCTATTAACATTCTGTAGGGGC 3'	5' primer for amplifying human kappa light chain constant region
(39)	5' CCGGAATTCTTATCATTTACCCGGAGA 3'	3' primer located in the CH3 region of human IgG1 to amplify the entire heavy chain
(40)	5' TCTGCACTAGTTGGAATGGGCACATGCAG 3'	3' primer for amplifying the Fd region of mouse IgM

The 19 primers listed in Table 5 have been listed in the Sequence Listing and have been assigned the following SEQ. ID. NO.:

(1)=SEQ. ID. NO. 40
(2)=SEQ. ID. NO. 41
(3)=SEQ. ID. NO. 42
(4)=SEQ. ID. NO. 43
(5)=SEQ. ID. NO. 44
(6)=SEQ. ID. NO. 45
(7)=SEQ. ID. NO. 46
(8)=SEQ. ID. NO. 47
(9)=SEQ. ID. NO. 48
(10)=SEQ. ID. NO. 49
(11)=SEQ. ID. NO. 50
(12)=SEQ. ID. NO. 51
(12A)=SEQ. ID. NO. 52
(13)=SEQ. ID. NO. 53
(14)=SEQ. ID. NO. 54
(15)=SEQ. ID. NO. 55
(16)=SEQ. ID. NO. 56
(17)=SEQ. ID. NO. 57
(18)=SEQ. ID. NO. 58

The 40 primers listed as "(1)" through "(40)" in Table 6 have also been individually and sequentially listed in the Sequence Listing beginning with SEQ. ID. NO. 59 through SEQ. ID. NO. 98, respectively.

b. Preparation of a Repertoire of Genes Encoding Immunoglobulin Variable Domain

Nitrophenylphosphonamidate (NPN) was selected as the ligand for receptor binding in preparing a heterodimeric receptor according to the methods of the invention.

Keyhole limpet hemocyanin (KLH) was conjugated to NPN to form a NPN-KLH conjugate used for immunizing a mouse to produce an anti-NPN immune response and thereby provide a source of ligand specific heterodimeric receptor genes.

The NPN-KLH conjugate was prepared by admixing 250 μl of a solution containing 2.5 mg of NPN in dimethylformamide with 750 μl of a solution containing 2 mg of KLH in 0.01 Molar (M) sodium phosphate buffer (pH 7.2). The two solutions were admixed by slow addition of the NPN solution to the KLH solution while the KLH solution was being agitated by a rotating stirring bar. Thereafter the admixture was maintained at 4° C. for 1 hour with the same agitation to allow conjugation to proceed. The conjugated NPN-KLH was isolated from the nonconjugated NPN and KLH by gel filtration through Sephadex G-25. The isolated NPN-KLH conjugate was injected into mice as described below.

The NPN-KLH conjugate was prepared for injection into mice by adding 100 μg of the conjugate to 250 μl of phosphate buffered saline (PBS). An equal volume of complete Freund's adjuvant was added and emulsified the entire solution for 5 minutes. A 129 G_{1x+} mouse was injected with 300 μl of the emulsion. Injections were given subcutaneously at several sites using a 21 gauge needle. A second immunization with NPN-KLH was given two weeks later. This injection was prepared as follows: 50 micrograms (μg) of NPN-KLH were diluted in 250 μl of PBS and an equal volume of alum was admixed to the NPN-KLH solution. The mouse was injected intraperitoneally with 500 μl of the solution using a 23 gauge needle. One month later the mice were given a final injection of 50 μg of the NPN-KLH conjugate diluted to 200 μl in PBS. This injection was given intravenously in the lateral tail vein using a 30 gauge needle. Five days after this final injection the mice were sacrificed and total cellular RNA was isolated from their spleens.

Total cellular RNA was prepared from the spleen of a single mouse immunized with KLH-NPN as described above using the RNA preparation methods described by Chomczynski et al., *Anal Biochem.*, 162:156–159 (1987) and using the RNA isolation kit (Stratagene) according to the manufacturer's instructions. Briefly, immediately after

removing the spleen from the immunized mouse, the tissue was homogenized in 10 ml of a denaturing solution containing 4.0M guanine isothiocyanate, 0.25M sodium citrate at pH 7.0, and 0.1M beta-mercaptoethanol using a glass homogenizer. One ml of sodium acetate at a concentration of 2M at pH 4.0 was admixed with the homogenized spleen. One ml of phenol that had been previously saturated with H_2O was also admixed to the denaturing solution containing the homogenized spleen. Two ml of a chloroform:isoamyl alcohol (24:1 v/v) mixture was added to this homogenate. The homogenate was mixed vigorously for ten seconds and maintained on ice for 15 minutes. The homogenate was then transferred to a thick-walled 50 ml polypropylene centrifuged tube (Fisher Scientific Company, Pittsburg, Pa.). The solution was centrifuged at 10.000×g for 20 minutes at 4° C. The upper RNA-containing aqueous layer was transferred to a fresh 50 ml polypropylene centrifuge tube and mixed with an equal volume of isopropyl alcohol. This solution was maintained at −20° C. for at least one hour to precipitate the RNA. The solution containing the precipitated RNA was centrifuged at 10.000×g for twenty minutes at 4° C. The pelleted total cellular RNA was collected and dissolved in 3 ml of the denaturing solution described above. Three ml of isopropyl alcohol was added to the re-suspended total cellular RNA and vigorously mixed. This solution was maintained at −20° C. for at least 1 hour to precipitate the RNA. The solution containing the precipitated RNA was centrifuged at 10.000×g for ten minutes at 4° C. The pelleted RNA was washed once with a solution containing 75% ethanol. The pelleted RNA was dried under vacuum for 15 minutes and then re-suspended in dimethyl pyrocarbonate (DEPC) treated (DEPC-H_2O) H_2O.

Messenger RNA (mRNA) enriched for sequences containing long poly A tracts was prepared from the total cellular RNA using methods described in *Molecular Cloning: A Laboratory Manual*, Maniatis et al., eds., Cold Spring Harbor, N.Y., (1982). Briefly, one half of the total RNA isolated from a single immunized mouse spleen prepared as described above was re-suspended in one ml of DEPC-H_2O and maintained at 65° C. for five minutes. One ml of 2×high salt loading buffer consisting of 100 mM Tris-HCl (Tris [hydroxymethyl] amino methane hydrochloride), 1M sodium chloride (NaCl), 2.0 mM disodium ethylene diamine tetra-acetic acid (EDTA) at pH 7.5, and 0.2% sodium dodecyl sulfate (SDS) was added to the re-suspended RNA and the mixture allowed to cool to room temperature. The mixture was then applied to an oligo-dT (Collaborative Research Type 2 or Type 3) column that was previously prepared by washing the oligo-dT with a solution containing 0.1M sodium hydroxide and 5 mM EDTA and then equilibrating the column with DEPC-H_2O. The eluate was collected in a sterile polypropylene tube and reapplied to the same column after heating the eluate for 5 minutes at 65° C. The oligo dT column was then washed with 2 ml of high salt loading buffer consisting of 50 mM Tris-HCl, pH 7.5, 500 mM sodium chloride, 1 mM EDTA at pH 7.5 and 0.1% SDS. The oligo dT column was then washed with 2 ml of 1×medium salt buffer consisting of 50 mM Tris-HCl, pH 7.5, 100 mM, 1 mM EDTA and 0.1% SDS. The messenger RNA was eluted from the oligo dT column with 1 ml of buffer consisting of 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, at pH 7.5, and 0.05% SDS. The messenger RNA was purified by extracting this solution with phenol/chloroform followed by a single extraction with 100% chloroform. The messenger RNA was concentrated by ethanol precipitation and re-suspended in DEPC H_2O.

The messenger RNA (mRNA) isolated by the above process contains a plurality of different V_H coding polynucleotides, i.e., greater than about 10^4 different V_H-coding genes, and contains a similar number of V_L-coding genes. Thus, the mRNA population represents a repertoire of variable region-coding genes.

c. Preparation of DNA Homologs

In preparation for PCR amplification, mRNA prepared above is used as a template for cDNA synthesis by a primer extension reaction. In a typical 50 µl transcription reaction, 5–10 ug of spleen mRNA in water is first hybridized (annealed) with 500 ng (50.0 pmol) of the 3' V_H primer (primer 12A, Table 5), at 65° C. for five minutes. Subsequently, the mixture is adjusted to 1.5 mM dATP, dCTP, dGTP and dTTP, 40 mM Tris-HCl, pH 8.0, 8 mM $MgCl_2$, 50 mM NaCl, and 2 mM spermidine. Moloney-Murine Leukemia virus Reverse transcriptase (Stratagene), 26 units, is added and the solution is maintained for 1 hour at 37 C.

PCR amplification is performed in a 100 µl reaction containing the products of the reverse transcription reaction (approximately 5 ug of the cDNA/RNA hybrid), 300 ng of 3' V_H primer (primer 12A of Table 5), 300 ng each of the 5' V_H primers (primers 2–10 of Table 5) 200 mM of a mixture of dNTP's, 50 mM KCl, 10 mM Tris-HCl pH 8.3, 15 mM $MgCl_2$, 0.1% gelatin and 2 units of *Thermus aquaticus* (Taq) DNA polymerase. The reaction mixture is overlaid with mineral oil and subjected to 40 cycles of amplification. Each amplification cycle includes denaturation at 92° C. for 1 minute, annealing at 52° C. for 2 minutes and polynucleotide synthesis by primer extension (elongation) at 72° C. for 1.5 minutes. The amplified V_H-coding DNA homolog containing samples are then extracted twice with phenol/chloroform, once with chloroform, ethanol precipitated and are stored at −70° C. in 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA.

Using unique 5' primers (2–9, Table 5), efficient V_H-coding DNA homolog synthesis and amplification from the spleen mRNA is achieved as shown by agarose gel electrophoresis. The amplified cDNA (V_H-coding DNA homolog) was seen as a major band of the expected size (360 bp). The amount the amplified V_H-coding polynucleotide fragment in each reaction is similar, indicating that all of these primers were about equally efficient in initiating amplification. The yield and quality of the amplification with these primers is reproducible.

The primer containing inosine also synthesizes amplified V_H-coding DNA homologs from spleen mRNA reproducibly, leading to the production of the expected sized fragment, of an intensity similar to that of the other amplified cDNAs. The presence of inosine also permits efficient DNA homolog synthesis and amplification, clearly indicating that such primers are useful in generating a plurality of V_H-coding DNA homologs. Amplification products obtained from the constant region primers (primers 11 and 13, Table 5) are more intense indicating that amplification was more efficient, possibly because of a higher degree of homology between the template and primers. Following the above procedures, a V_H-coding gene library is constructed from the products of eight amplifications, each performed with a different 5' primer. Equal portions of the products from each primer extension reaction are mixed and the mixed product is then used to generate a library of V_H-coding DNA homolog-containing vectors.

DNA homologs of the V_L are also prepared from the purified mRNA prepared as described above. In preparation for PCR amplification, mRNA prepared according to the above examples is used as a template for cDNA synthesis. In a typical 50 µl transcription reaction, 5–10 ug of spleen

mRNA in water is first annealed with 300 ng (50.0 pmol) of the 3' V_L primer (primer 14, Table 5), at 65° C. for five minutes. Subsequently, the mixture is adjusted to 1.5 mM dATP, dCTP, dGTP, and dTTP, 40 mM Tris-HCl, pH 8.0, 8 mM $MgCl_2$, 50 mM NaCl, and 2 mM spermidine. Moloney-Murine Leukemia virus reverse transcriptase (Stratagene), 26 units, is added and the solution is maintained for 1 hour at 37° C. The PCR amplification is performed in a 100 µl reaction containing approximately 5 ug of the cDNA/RNA hybrid produced as described above, 300 ng of the 3' V_L primer (primer 14 of Table 5), 300 ng of the 5' V_L primer (primer 16 of Table 5), 200 mM of a mixture of dNTP's, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 15 mM $MgCl_2$, 0.1% gelatin and 2 units of Taq DNA polymerase. The reaction mixture is overlaid with mineral oil and subjected to 40 cycles of amplification. Each amplification cycle includes denaturation at 92° C. for 1 minute, annealing at 52° C. for 2 minutes and elongation at 72° C. for 1.5 minutes. The amplified samples are then extracted twice with phenol/chloroform, once with chloroform, ethanol precipitated and are stored at −70° C. in 10 mM Tris-HCl, 7.5 and 1 mM EDTA.

d. Insertion of DNA Homologs into a DNA Expression Vector

To prepare an expression library enriched in V_H sequences, DNA homologs enriched in V_H sequences are prepared according to Example 2c using the same set of 5' primers but with primer 12A (Table 5) as the 3' primer. The resulting PCR amplified products (2.5 ug/30 µl of 150 mM NaCl, 8 mM Tris-HCl, pH 7.5, 6 mM $MgSO_4$, 1 mM DTT, 200 ug/ml BSA) are digested at 37° C. with restriction enzymes Xho I (125 units) and Spe I (125 units). In cloning experiments which required a mixture of the products of the amplification reactions, equal volumes (50 µl, 1–10 ug concentration) of each reaction mixture are combined after amplification but before restriction digestion. The V_H homologs are purified on a 1% agarose gel using the standard electro-elution technique described in *Molecular Cloning A Laboratory Manual*, Maniatis et al., eds., Cold Spring Harbor, N.Y., (1982). After gel electrophoresis of the digested PCR amplified spleen mRNA, the region of the gel containing DNA fragments of approximate 350 bps is excised, electro-eluted into a dialysis membrane, ethanol precipitated and re-suspended suspended in a TE solution containing 10 mM Tris-HCl, pH 7.5 and 1 mM EDTA to a final concentration of 50 ng/µl. The resulting V_H DNA homologs represent a repertoire of polypeptide genes having cohesive termini adapted for directional ligation to the vector Lambda Hc2. These prepared V_H DNA homologs are then directly inserted by directional ligation into linearized Lambda Hc2 expression vector prepared as described below.

The Lambda Hc2 expression DNA vector is prepared for inserting a DNA homolog by admixing 100 ug of this DNA to a solution containing 250 units each of the restriction endonucleases Xho I and Spe I (both from Boehringer Mannheim, Indianapolis, Ind.) and a buffer recommended by the manufacturer. This solution is maintained at 37 from 1.5 hours. The solution is heated at 65° C. for 15 minutes top inactivate the restriction endonucleases. The solution is chilled to 30° C. and 25 units of heat-killable (HK) phosphatase (Epicenter, Madison, Wis.) and $CaCl_2$ is admixed to it according to the manufacturer's specifications. This solution is maintained at 30° C. for 1 hour. The DNA is purified by extracting the solution with a mixture of phenol and chloroform followed by ethanol precipitation. The Lambda Hc2 expression vector is now ready for ligation to the V_H DNA homologs prepared in the above examples. These prepared V_H DNA homologs are then directly inserted into the Xho I and Spe I restriction digested Lambda Hc2 expression vector that prepared above by ligating 3 moles of V_H DNA homolog inserts with each mole of the Hc2 expression vector overnight at 5° C. Approximately 3.0×10^5 plague forming units are obtained after packaging the DNA with Gigapack II Bold (Stratagene) of which 50% are recombinants. The ligation mixture containing the V_H DNA homologs are packaged according to the manufacturers specifications using Gigapack Gold II Packing Extract (Stratagene). The resulting Lambda Hc2 expression libraries are then transformed into XL1-Blue cells.

To prepare a library enriched in V_L sequences, PCR amplified products enriched in V_L sequences are prepared according to Example 2c. These V_L DNA homologs are digested with restriction enzymes Sac I and Xba I and the digested V_L DNA homologs are purified on a 1% agarose gel as described above for the V_H DNA homologs to form a repertoire of V_L-polypeptide genes adapted for directional ligation. The prepared V_L DNA homologs are then directionally ligated into the Lambda Lc2 expression vector previously digested with the restriction enzymes, Sac I and Xba I as described for Lambda Hc2. The ligation mixture containing the V_L DNA homologs is packaged to form a Lambda Lc2 expression library as described above and is ready to be plated on XL1-Blue cells.

e. Randomly Combining V_H and V_L DNA Homologs on the Same Expression Vector

The construction of a library containing vectors for expressing two cistrons that express heavy and light chains is accomplished in two steps. In the first step, separate heavy and light chain libraries are constructed in the expression vectors Lambda Hc2 and Lambda Lc2, respectively, as described using gene repertoires obtained from a mouse immunized with NPN-KLH. In the second step, these two libraries are combined at the antisymmetric EcoR I sites present in each vector. This resulted in a library of clones each of which potentially co-expresses a heavy and a light chain. The actual combinations are random and do not necessarily reflect the combinations present in the B-cell population in the parent animal.

The spleen mRNA resulting from the above immunizations (Example 2b) is isolated and used to create a primary library of V_H gene sequences using the Lambda Hc2 expression vector. The primary library contains 1.3×10^6 plaque-forming units (pfu) and can be screened for the expression of the decapeptide tag to determine the percentage of clones expressing V_H and C_H1 (Fd) sequences. The sequence for this peptide is only in frame for expression following the cloning of a Fd (or V_H) fragment into the vector. At least 80% of the clones in the library express Fd fragments based on immunodetection of the decapeptide tag.

The light chain library is constructed in the same way as the heavy chain and contains 2.5×10^6 members. Plaque screening, using an anti-kappa chain antibody, indicates that 60% of the library contained express light chain inserts. A small percentage of inserts results from incomplete dephosphorylation of vector after cleavage with Sac I and Xba I.

Once obtained, the two libraries are used to construct a combinatorial library by crossing them at the EcoR I site. To accomplish the cross, DNA is first purified from each library.

The Lambda Lc2 library prepared in Example 2d is amplified and 500 ug of Lambda Lc2 expression library phage DNA is prepared from the amplified phage stock using the procedures described in *Molecular Cloning: A Laboratory Manual*, Maniatis et al., eds., Cold Spring Harbor, N.Y. (1982). Fifty ug of this amplified expression

library phage DNA is maintained in a solution containing 100 units of MLu I restriction endonuclease (Boehringer Mannheim. Indianapolis. Ind.) in 200 µl of a buffer supplied by the endonuclease manufacturer for 1.5 hours at 37° C. The solution is then extracted with a mixture of phenol and chloroform. The DNA is then ethanol precipitated and re-suspended in 100 µl of water. This solution is admixed with 100 units of the restriction endonuclease EcoR I (Boehringer) in a final volume of 200 µl of buffer containing the components specified by the manufacturer. This solution is maintained at 37° C. for 1.5 hours and the solution is then extracted with a mixture of phenol and chloroform. The DNA was ethanol precipitated and the DNA re-suspended in TE.

The Lambda Hc2 expression library prepared in Example 2d is amplified and 500 ug of Lambda Hc2 expression library phage DNA is prepared using the methods detailed above. 50 ug of this amplified library phage DNA is maintained in a solution containing 100 units of Hind III restriction endonuclease (Boehringer) in 200 µl of a buffer supplied by the endonuclease manufacturer for 1.5 hours at 37° C. The solution is then extracted with a mixture of phenol and chloroform saturated with 0.1M Tris-HCl. pH 7.5. The DNA is then ethanol precipitated and re-suspended in 100 µl of water. This solution is admixed with 100 units of the restriction endonuclease EcoR I (Boehringer) in a final volume of 200 µl of buffer containing the components specified by the manufacturer. This solution is maintained at 37° C. for 1.5 hours and the solution is then extracted with a mixture of phenol and chloroform. The DNA is ethanol precipitated and the DNA re-suspended in TE.

The restriction digested Hc2 and Lc2 expression libraries are ligated together. To that end. a DNA admixture consists of 1 ug of Hc2 and 1 ug of Lc2 phage library DNA is prepared in a 10 µl reaction using the reagents supplied in a ligation kit (Stratagene). The DNA admixture is warmed to 45° C. for 5 minutes to melt any cohesive termini that may have reannealed. The admixture is then chilled to 0° C. to prevent religation. Bacteriophage T4 DNA ligase (0.1 Weiss units which is equivalent to 0.02 units as determined in an exonuclease resistance assay) is admixed into the chilled DNA solution along with 1 ul of 5 mM ATP and 1 ul 10×bacteriophage T4 DNA ligase buffer (10×buffer is prepared by admixing 200 mM Tris-HCl. pH 7.6. 50 mM $Mgcl_2$.50 mM DTT. and 500 ug/ml BSA) to form a ligation admixture. After ligation for 16 hr at 4° C.. 1 µl of the ligated the phage DNA is packaged with Gigapack Gold II packaging extract and plated on XL1-Blue cells prepared according to the manufacturers instructions to form a Lambda phage library of dicistronic expression vectors capable of expressing heavy and light chains derived from the NPN-immunized mouse. A portion of the clones obtained are used to determine the effectiveness of the combination.

f. Selection of Anti-NPN Reactive Heterodimer-Producing Dicistronic Vectors

The combinatorial Fab expression library prepared above in Example 2a was screened to identify clones having affinity for NPN. To determine the frequency of the phage clones which co-expressed the light and heavy chain fragments. duplicate lifts of the light chain. heavy chain and combinatorial libraries were screened as above for light and heavy chain expression. In this study of approximately 500 recombinant phage. approximately 60% co-expressed light and heavy chain proteins.

All three libraries. the light chain. the heavy chain and the combinatorial. were screened to determine if they contained recombinant phage that expressed antibody fragments which bound NPN. In a typical procedure 30.000 phage were plated on XL1-Blue cells and duplicate lifts with nitrocellulose were screened for binding to NPN coupled to ^{125}I labeled BSA. The BSA was iodinated following the Chloramine-T method as described by Bolton et al.. *Biochem.*, 133:529–534 (1973). Duplicate screens of 80.000 recombinant phage from the light chain library and a similar number from the heavy chain library did not identify any clones which bound the antigen. In contrast. the screen of a similar number of clones from the Fab expression library identified many phage plaques that bound NPN. This observation indicates that under conditions where many heavy chains in combination with light chains bind to antigen the same heavy or light chains alone do not. Therefore. in the case of NPN. it is believed that there are many heavy and light chains that only bind antigen when they are combined with specific light and heavy chains respectively.

To assess the ability to screen large numbers of clones and obtain a more quantitative estimate of the frequency of antigen binding clones in the combinatorial library. one million phage plaques were screened and approximately 100 clones which bound to antigen were identified. For six clones which were believed to bind NPN. a region of the plate containing the six positive and approximately 20 surrounding bacteriophage plaques was selected and each plaque was cored. replated. and screened with duplicate lifts. As expected. approximately one in twenty of the phage specifically bound to antigen. Cores of regions of the plated phage believed to be negative did not give positives on replating.

Clone 2b. one of the plaques which reacted with NPN. was excised according to an in vivo excision protocol where 200 ul of phage stock and 200 ul of a F+ derivative of XL1-Blue (A_{600}=1.00) (Stratagene) were admixed with 1 ul of M13mp8 helper phage (1×10^{10} pfu/milliliter (ml)) and maintained at 37° C. for 15 minutes. After a four hour maintenance in Luria-Bertani (LB) medium and heating at 70° C. for 20 minutes to heat kill the XL1-Blue cells. the phagemids were re-infected into XL1-Blue cells and plated onto LB plates containing ampicillin. This procedure converted the cloned insert from the Lambda Zap II vector into a plasmid vector to allow easy manipulation and sequencing (Stratagene). The phagemid DNA encoding the V_H and part of the V_L was then determined by DNA sequencing using the Sanger dideoxy method described in Sanger et al.. *Proc. Natl. Acad. Sci.*, 74:5463–5467 (1977) using a Sequenase kit according to manufacturer's instructions (US Biochemical Corp.. Cleveland. Ohio). The nucleotide residue sequence of Clone 2b Fd chain is listed in the Sequence Listing as SEQ. ID. NO. 99. The nucleotide residue sequences of the kappa light chain variable and constant regions are listed in the Sequence Listing as SEQ. ID. NO. 100 and SEQ. ID. NO. 101. respectively.

g. Preparation of a DNA Sequence Encoding a Filamentous Phage Coat Protein Membrane Anchor

cpVIII Membrane Anchor: M13mp18. a commercially available bacteriophage vector (Pharmacia. Piscataway. N.J.). was used as a source for isolating the gene encoding cpVIII. The sequence of the gene encoding the membrane anchor domain of cpVIII listed in Sequence Listing as SEQ. ID. NO. 102. was modified through PCR amplification to incorporate the restriction endonuclease sites. Spe I and EcoR I. and two stop codons prior to the EcoR I site. The corresponding amino acid residue sequence of the membrane anchor domain of cpVIII is listed as SEQ. ID. NO. 17.

To prepare a modified cpVIII. replicative form DNA from M13mp18 was first isolated. Briefly. into 2 ml of LB

(Luria-Bertani medium). 50 ul of a culture of a bacterial strain carrying an F' episome (JM107, JM109 or TG1) was admixed with a one tenth suspension of bacteriophage particles derived from a single plaque. The admixture was incubated for 4 to 5 hours at 37° C. with constant agitation. The admixture was then centrifuged at 12,000×g for 5 minutes to pellet the infected bacteria. After the supernatant was removed, the pellet was resuspended by vigorous vortexing in 100 ul of ice-cold solution I. Solution I was prepared by admixing 50 mM glucose, 10 mM EDTA and 25 mM Tris-HCl, pH 8.0, and autoclaving for 15 minutes.

To the bacterial suspension, 200 ul of freshly prepared Solution II was admixed and the tube was rapidly inverted five times. Solution II was prepared by admixing 0.2N NaOH and 1% SDS. To the bacterial suspension, 150 ul of ice-cold Solution III was admixed and the tube was vortexed gently in an inverted position for 10 seconds for to disperse Solution III through the viscous bacterial lysate. Solution III was prepared by admixing 60 ml of 5M potassium acetate, 11.5 ml of glacial acetic acid and 28.5 ml of water. The resultant bacterial lysate was then stored on ice for 5 minutes followed by centrifugation at 12,000×g for 5 minutes at 4° C. in a microfuge. The resultant supernatant was recovered and transferred to a new tube. To the supernatant was added an equal volume of phenol: chloroform and the admixture was vortexed. The admixture was then centrifuged at 12,000×g for 2 minutes in a microfuge. The resultant supernatant was transferred to a new tube and the double-stranded bacteriophage DNA was precipitated with 2 volumes of ethanol at room temperature. After allowing the admixture to stand at room temperature for 2 minutes, the admixture was centrifuged to pellet the DNA. The supernatant was removed and the pelleted replicative form DNA was resuspended in 25 ul of Tris-HCl, pH 7.6, and 10 mM EDTA (TE).

The double-stranded M13mp18 replicative form DNA was then used as a template for PCR. Primers, AK 5 (SEQ. ID. NO. 103) and AK 6 (SEQ. ID. NO. 104), the sequences of which are listed in Table 7 below, were used in the PCR reaction to amplify the mature gene for cpVIII member anchor domain and incorporate the two cloning sites, Spe I and EcoR I. For the PCR reaction, 2 ul containing 1 nanogram (ng) of M13mp18 replicative form DNA was admixed with 10 ul of 10×PCR buffer purchased commercially (Promega Biotech, Madison, Wis.) in a 0.5 ml microfuge tube. To the DNA admixture, 8 ul of a 2.5 mM solution of dNTPs (dATP, dCTP, dGTP, dTTP) was admixed to result in a final concentration of 200 micromolar (uM). Three ul (equivalent to 60 picomoles (pM)) of the 5' forward AK 5 primer and 3 ul (60 pM) of the 3' backward AK 6 primer was admixed into the DNA solution. To the admixture, 73 ul of sterile water and 1 ul/5 units of polymerase (Promega Biotech) was added. Two drops of mineral oil were placed on top of the admixture and 40 rounds of PCR amplification in a thermocycler were performed. The amplification cycle consisted of 52° C. for 2 minutes, 72° C. for 1.5 minutes and 91° C. for 2 minutes. The resultant PCR modified cpVIII membrane anchor domain DNA fragment from M13mp18 containing samples were then purified with Gene Clean (BIO101, La Jolla, Calif.), extracted twice with phenol/chloroform, once with chloroform followed by ethanol precipitation and were stored at −70° C. in 10 mM Tris-HCl, pH 7.5, and 1 mM EDTA.

TABLE 7

SEQ. ID NO.	Primer	
(103)[1]	AK 5 (F)	5' GTGCCCAGGGATTGTACTAGTGCTGAGGGTGACGAT 3'
(104)[2]	AK 6 (B)	5' ACTCGAATTCTATCAGCTTGCTTTCGAGGTGAA 3'
(105)[3]	Hc3 (F)	5' AGGTCCAGCTTCTCGAGTCTGG 3'
(106)[4]	AK 7 (B)	5' GTCACCCTCAGCACTAGTACAATCCCTGGGCAC 3'
(107)[5]	G-3 (F)	5' GAGACGACTAGTGGTGGCGGTGGCTCTCCATTC GTTTGTGAATATCAA 3'
(108)[6]	G-3 (B)	5' TTACTAGCTAGCATAATAACGGAATACCCAAAA GAACTGG 3'
(109)[7]	LAC-F	5' TATGCTAGCTAGTAACACGACAGGTTTCCCGAC TGG 3'
(110)[8]	LAC-B	5' ACCGAGCTCGAATTCGTAATCATGGTC 3'

F Forward Primer
B Beckward Primer
[1]From 5' to 3': the overlapping sequence for C_H1 3' end is double underlined; the Spe I restriction site sequence is single underlined; the overlapping sequence for cpVIII is double underlined.
[2]EcoR I restriction site sequence is single underlined
[3]Xho I restriction site sequence is underlined
[4]From 5' to 3': the overlapping sequence for cpVIII is double underlined; the Spe I restriction site sequence is single underlined; the overlapping sequence for C_H1 3' end is double underlined.
[5]From 5' to 3': Spe I restriction site sequence is single underlined; the overlapping sequence with the 5' end of cpIII is double underlined
[6]From 5' to 3': Nhe I restriction site sequence is single underlined; the overlapping sequence with 3' end of cpIII is double underlined.
[7]From 5' to 3': overlapping sequence with the 3' end of cpIII is double underlined; Nhe I restriction sequence begins with the nucleotide residue "G" at position 4 and extends 5 more residues = GCTAGC.
[8]EcoR I restriction site sequence is single underlined.

To verify amplification of the modified cpVIII membrane anchor domain, the PCR purified DNA products were electrophoresed in a 1% agarose gel. The expected size of the cpVIII was approximately 150 base pairs. The area in the agarose containing the modified cpVIII DNA fragment was isolated from the agarose as described above. The sequence of the isolated modified cpVIII DNA fragment is listed as SEQ. ID. NO. 111. The isolated cpVIII DNA fragment was then admixed with a similarly prepared fragment of modi-

fied Fd as described below in Example 2i in order to form a DNA segment encoding the fusion protein Fd-cpVIII.

cpIII Membrane Anchor: M13mp18 was also used as a source for isolating the gene encoding the membrane anchor domain at cpIII, the sequence of which is listed in the Sequence Listing as SEQ. ID. NO. 112. The amino acid residue sequence of membrane anchor domain cpIII is listed in SEQ. II). NO. 16. M13mp18 replicative form DNA was prepared as described above and used as a template for PCR for amplifying the mature gene for cpIII membrane anchor domain and incorporating the two cloning sites, Spe I and EcoR I.

The primer pair, G-3(F) (SEQ. ID. NO. 107) and G-3(B) (SEQ. ID. NO. 108) listed in Table 7, was used in PCR as performed above to incorporate Spe I and Nhe I restriction sites. The resultant PCR modified cpIII DNA fragment was verified and purified as described above. The sequence of the PCR modified cpIII membrane anchor domain DNA fragment is listed in the Sequence Listing as SEQ. ID. NO. 113. A second PCR amplification using the primer pairs, Lac-F (SEQ. ID. NO. 109) and Lac-B (SEQ. ID. NO. 110) listed in Table 7, was performed on a separate aliquot of M13mp18 replicative form template DNA. The primers used for this amplification were designed to incorporate an overlapping sequence with the nucleotides encoding the membrane anchor region of cpIII and the adjacent Nhe I site along with a sequence encoding a LacZ promoter region 5' to an EcoR I restriction site. The reaction and purification of the PCR product performed as described above. The sequence of the resultant PCR modified cpIII DNA fragment having Nho I and EcoR I restriction sites is listed in the Sequence Listing as SEQ. ID. NO. 114.

The products of the first and second PCR amplifications were then recombined at the nucleotides corresponding to cpIII membrane anchor overlap and Nhe I restriction site and subjected to a second round of PCR using the G3-F (SEQ. ID. NO. 107) and Lac-B (SEQ. ID. NO. 110) primer pair to form a recombined PCR DNA fragment product consisting of the following: a 5' Spe I restriction site; a cpIII DNA membrane anchor domain beginning at the nucleotide residue sequence which corresponds to the amino acid residue 198 of the entire mature cpIII protein; an endogenous stop site provided by the membrane anchor at amino acid residue number 112; a Nhe I restriction site, a LacZ promoter sequence; and a 3' EcoR I restriction site. The recombined PCR modified cpIII membrane anchor domain DNA fragment was then restriction digested with Spe I and EcoR I to produce a DNA fragment for directional ligation into a pComb phagemid expression vector prepared in Example 1a(iv) and to form a pCombIII phagemid expression vector as described in Example 1b(ii).

h. Isolation of Anti-NPN Coding V_H DNA Segment

To prepare modified Fd fragments for recombination with the PCR modified cpVIII membrane anchor domain fragment to form a Fd-cpVIII DNA fusion product, PCR amplification as described above was performed using Clone 2b, prepared in Example 2f, as a template. The primers, Hc3 (SEQ. ID. NO. 105) and AK 7 (SEQ. ID. NO. 106), the sequences of which are listed in Table 7, were used in PCR to amplify the Fd portion of the Clone 2b and incorporate Xho I and Spe I cloning sites along with a cpVIII overlapping sequence. The amplified PCR modified Fd product was purified, electrophoresed and isolated from 1% agarose gels as described above. The size of the Fd fragment was 680 base pairs.

i. Preparation of a DNA Segment Encoding a Portion of the Fusion Protein Fd-cpVIII

The purified PCR modified Fd DNA fragment containing cpVIII overlapping nucleotide sequences prepared above was then admixed with the PCR modified cpVIII membrane anchor domain fragment to form an admixture. The fragments in the admixture were allowed to recombine at their complementary regions. The admixture containing the recombined PCR fragments was then subjected to a second round of PCR amplification as described above using the end primer pair AK 6 (SEQ. ID. NO. 104) and Hc3 (SEQ. ID. NO. 105) (Table 7). The corresponding product of the PCR amplification was purified and electrophoresed on agarose gels as described above. The PCR product was determined to be approximately 830 base pairs (Fd=680+150) confirming the fusion of Fd with cpVIII. The sequence of the PCR product linking the Fd sequence with the cpVIII sequence in frame in a 5' to 3' direction is listed as SEQ. ID. NO. 115. The Fd-cpVIII fusion product was then used in directional ligations described in Example 2j for the construction of a pCBAK8-2b dicistronic phagemid expression vector.

j. Construction of pCBAK8-2b Dicistronic Expression Vector

To construct a phagemid vector for the coordinate expression of a Fd-cpVIII fusion protein with kappa light chain, the PCR amplified Fd-cpVIII fusion product prepared in above in Example 2i was first ligated into Clone 2b phagemid expression vector isolated from the NPN combinatorial library prepared in Example 2f. For the ligation, the Fd-cpVIII PCR fusion product was first restriction digested with Xho I and EcoR I. Clone 2b phagemid vector was similarly digested resulting in the removal of the cloning and decapeptide regions. The digested Fd-cpVIII fragment was admixed and ligated into the digested Clone 2b at the cohesive termini generated by Xho I and EcoR I restriction digestion. The ligation resulted in operatively linking the nucleotide residue sequence encoding the Fd-cpVIII polypeptide fusion protein to a second cassette having the nucleotide residue sequences encoding the ribosome binding site, a pelB leader sequence and the kappa light chain already present in Clone 2b to form a dicistronic DNA molecule in the original Clone 2b phagemid expression vector.

E. coli, strain TG1, was then transformed with the phagemid containing the dicistronic DNA molecule and transformants were selected on ampicillin as the original Clone 2b contained an ampicillin selectable resistance marker gene. For high efficiency electro-transformation of *E. coli*, a 1:100 volume of an overnight culture of TG1 cells was inoculated into one liter of L-broth (1% Bacto tryptone, 0.5% Bacto yeast extract, 0.5% NaCl). The cell suspension was maintained at 37° C. with vigorous shaking to a absorbance at 600 nm of 0.5 to 1.0. The cell suspension in log phase growth was then harvested by first chilling the flask on ice for 15 to 30 minutes followed by centrifugation in a cold rotor at 4000 g for 15 minutes to pellet the bacteria. The resultant supernatant was removed and the bacterial cell pellet was resuspended in a total of one liter of cold water to form a cell suspension. The centrifugation and resuspension procedure was repeated two more times and after the final centrifugation, the cell pellet was resuspended in 20 ml of cold 10% glycerol. The resuspended cell suspension was then centrifuged to form a cell pellet. The resultant cell pellet was resuspended to a final volume of 2 to 3 ml in cold 10% glycerol resulting in a cell concentration of 1 to 3×10^{10} cells/ml. For the electro-transformation procedure, 40 ul of the prepared cell suspension was admixed with 1 to 2 ul of phagemid DNA to form a cell-phagemid DNA admixture.

The resultant admixture was mixed and allowed to sit on ice for one minute. An electroporation apparatus. for example a Gene Pulsar. was set a 25 uF and 2.5 kV. The pulse controller was set to 200 ohms. The cell-DNA admixture was transferred to a cold 0.2 cm electroporation cuvette. The cuvette was then placed in the chilled safety chamber and pulsed once at the above settings. To the pulsed admixture. 1 ml of SOC medium was then admixed and the cells were resuspended with a Pasteur pipette (SOC medium was prepared by admixing 2% Bacto tryptone. 0.5% Bacto yeast extract. 10 mM NaCl. 2.5 mM KCl. 10 mM $MgCl_2$. 10 mM $MgSO_4$. and 20 mM glucose). The cells suspension was then transferred to a 17×100 mm polypropylene tube and maintained at 37° C. for one hour. After the maintenance period. the transformed TG1 cells were then plated on ampicillin LB plates for selection of ampicillin resistant colonies containing the phagemid which provided the selectable marker gene.

Ampicillin resistant colonies were selected and analyzed for the correct insert size and expression of Fab. Briefly. DNA minipreps of selected colonies were prepared for the isolation of phagemid DNA. The isolated phagemid DNA from each miniprep was restriction digested with Xho I and EcoR I and the digests were electrophoresed on a 1% agarose gel. Clone AK16 was selected as an 830 bp fragment was visualized on the gels confirming the insertion of the Fd-cpVIII PCR fusion product into digested Clone 2b.

Clone AK16 phagemid was then restriction digested with Xho I and Xba I and the nucleotide residue sequence of the dicistronic DNA molecule encoding the Fd-cpVIII fusion protein. the ribosome binding site and pelB leader sequence for expression of the light chain. a spacer region and the 2b kappa light chain was isolated by agarose gel electrophoresis. The isolated dicistronic DNA fragment was then ligated into a Xho I and Xba I restriction digested pCBAKO expression vector prepared in Example 1c(ii) to form a dicistronic phagemid expression vector designated pCBAK8-2b.

The resultant pCBAK8-2b expression vector consisted of nucleotide residue sequences encoding the following elements: f1 filamentous phage origin of replication; a chloramphenicol acetyl transferase selectable resistance marker gene; an inducible LacZ promoter upstream from the LacZ gene; a multiple cloning site flanked by T3 and T7 polymerase promoters; and the dicistronic DNA molecule (a first cassette consisting of a ribosome binding site. a pelB leader. and a Fd-cpVIII DNA fusion product operatively linked to a second cassette consisting of a second ribosome binding site. a second pelb leader. and a kappa light chain).

k. Construction of pCBAK3-2b Dicistronic Expression Vector

To construct a phagemid vector for the coordinate expression of a Fd-cpIII fusion protein with kappa light chain. the PCR amplified and recombined cpIII membrane anchor prepared in Example 2g having a 5' Spe I and 3' EcoR I restriction site was first directionally ligated into a pComb phagemid expression vector prepared in Example 1a(iv) to form a pCombIII phagemid vector. See Example 1b(ii) for details of vector construction. The resultant pCombIII phagemid vector was then restriction digested with Sac II and Apa I to form an isolated fragment. The resultant isolated fragment containing the expression control sequences and the cpIII sequence was then directionally ligated into a similarly digested pCBAKO phagemid vector prepared in Example 1c(ii) to form a pCBAK3 phagemid expression vector. This vector lacked Fd and kappa light chain sequences.

A phagemid expression vector. pCBAK3-2b. for the expression of a fusion protein and kappa light chain was then constructed. Briefly. the pCBAK3 phagemid expression vector prepared above was first digested with Xho I and Spe I to form a linearized pCBAK3 phagemid expression vector. PCR amplified and modified Fd fragment. prepared in Example 2h containing Xho I and Spe I sites. was subsequently restriction digested with Xho I and Spe I. The resultant Fd fragment was then directionally ligated via cohesive termini into the Xho I and Spe I restriction digested pCBAK3 phagemid expression vector to form a second phagemid expression vector in which the PCR modified Fd fragment was operatively linked in-frame to nucleotide residue sequences encoding cpIII. *E. coli* strain XL1-Blue (Stratagene) was then transformed with the above phagemid vector containing Fd-cpIII. Transformants containing the Fd-cpIII encoding phagemid were selected on chloramphenicol. Phagemid DNA was isolated from chloramphenicol resistant clones and was restriction digested with Sac I and Xba I to form a linearized phagemid expression vector into which a Sac I and Xba I light chain fragment prepared below was directionally ligated.

Phagemid Clone 2b. isolated from the original combinatorial library as described in Example 2a. was restriction digested with Sac I and Xba I to isolate the nucleotide residue sequence encoding the kappa light chain. The isolated kappa light chain sequence was then directionally ligated into the Sac I and Xba I restriction digested phagemid expression vector prepared above containing Fd-cpIII to form the phagemid expression vector. pCBAK3-2b. The resultant vector contained the nucleotide residue sequence of a dicistronic DNA molecule for the coordinate expression of a Fd-cpIII fusion protein with kappa light chain. The resultant phagemid expression vector consisted of nucleotide residue sequences encoding the following elements: f1 filamentous phage origin of replication; a chloramphenicol acetyl transferase selectable resistance marker gene; an inducible LacZ promoter upstream from the LacZ gene; a multiple cloning site flanked by T3 and T7 polymerase promoters; and the dicistronic molecule (a first cassette consisting of a first ribosome binding site and pelB leader operatively linked to Fd-cpIII operatively linked to a second cassette consisting of a second LacZ. a second ribosome binding site. and a second pelB leader operatively linked to a kappa light chain).

XL1-Blue cells were then transformed with the phagemid expression vector pCBAK3-2b. Transformed colonies containing the chloramphenicol resistant phagemids were selected as described above and analyzed for the correct size insert and expression of Fab as described in Example 2j. Following verification of the insert and expression of Fab in the pCBAK3-2b phagemid vector. XL1-Blue cells were then transformed and induced for the expression of Fab antibodies as described in Examples 3 and 4.

3. Expression of Anti-NPN Heterodimer on Phage Surfaces

For expression of antibody Fab directed against NPN on phage surfaces. XL1-Blue cells were separately transformed with the phagemid vectors. pCBAK8-2b and pCBAK3-2b. prepared in Examples 2j and 2k. respectively. The transformants were selected on LB plates containing 30 ug/ml chloramphenicol. Antibiotic resistant colonies were selected for each phagemid transformation and grown in liquid cultures at 37° C. in super broth (super broth was prepared by admixing the following: 20 g 3 [N-Morpholino] propanesulfonic acid (MOPS); 60 g tryptone; 40 g yeast extract; and

2 liter of water; adjust pH to 7.0 with 10 m NaOH) containing 30 ug/ml chloramphenicol and 12.5 ug/ml tetracycline for the respective antibiotic selection of the phagemid and the F' episome. The antibiotic resistant transformed XL1-Blue cells were diluted to an optical density (OD_{600} nm) of 0.4 in super broth. The inducer, isopropyl thiogalactopyranoside (IPTG), was admixed to the bacterial suspension for a final concentration of 1 mM and the admixture was maintained at 37° C. for 1 hour to induce the expression of the fusion protein and kappa light chain from the LacZ promoter. Helper phage, either R408 or VCS M13 (Stratagene), was then admixed to the induced bacterial suspension at a ratio of 10–20 helper phage to 1 transformed bacterial cell to initiate the generation of copies of the sense strand of the phagemid DNA. The admixture containing the helper phage was then maintained for an additional two hours at 37° C. to allow for filamentous bacteriophage assembly wherein the expressed anti-NPN Fab antibodies fused to either bacteriophage membrane anchor domains of cpVIII or cpIII were incorporated into surface of the bacteriophage particles. The bacterial suspension was then centrifuged resulting in a bacterial cell pellet and a supernatant containing phage. The supernatant was removed, collected and assayed as described below for the presence of functional anti-NPN Fab molecules anchored to the phage particles by either cpVIII or cpIII.

4. Assays for Verifying the Presence and Function of Anti-NPN Heterodimer on the Surface of Filamentous Phage

a. Electron Microscopy

To localize functional Fab molecules, the binding to antigen labelled with colloidal gold was studied. Phage containing supernatants and bacterial cells prepared in Example 3 were spotted on formvar Polysciences, Inc., Warrington, Pa.) coated grids affixed onto a solid phase. In some experiments grids were coated with cells and infected with phage in situ. Subsequently grids were blocked with bovine serum albumin (BSA) 1% in PBS at pH 7.2, washed and incubated with 2–7 nanometer (nm) colloidal gold particles coated with BSA-NPN hapten conjugate for a time period sufficient to form a labeled immunoreaction complex. The grids were washed to remove excess gold particles and negatively stained in uranylacetate and visualized by electron microscopy.

Examination of filamentous phage and permeabilized cells producing phage revealed specific labelling of phage or exposed bacterial membranes. Phage were observed to contain 1 to 24 copies of antigen binding sites per particle. Neither helper phage alone nor intact *E. coli* labelled with antigen. Background nonspecific binding was very low. Filamentous phage particles emerging from the *E. coli* surfaces were labelled with antigen as shown in FIG. 9.

The generation of a related phage surface expression vector utilizing cpIII as a fusion partner with Clone 2b, pCBAK3-2b, revealed specific antigen labelling to the phage head but not the column. Additionally human anti-tetanus Fab expressed as a cpIII fusion did not bind to BSA-NPN antigen.

b. Phase Elisa

Microtitration plates were coated with NPN-BSA conjugate (0.1 ml, 1 μg/ml in 0.1M Tris-HCl pH 9.2), and blocked with 1% BSA in PBS. Serial two fold dilutions of pCBAK8-2b derived phage (0.1 ml), prepared in Example 3, were added to the pre-coated microtitration plate and incubated for 3 hours at ambient temperature or 16 hours at 4° C. The plates were washed with PBS and goat anti-kappa alkaline phosphatase conjugate (Fisher Biotech, Pittsburgh, Pa.) added (0.1 ml diluted 1/1000 in PBS containing 0.1% BSA) and incubated for 2 hours at room temperature. The plates were washed in PBS and substrate added (0.1 ml, 1 mg/ml p-nitrophenylphosphate in 0.1M Tris-HCl, pH 9.5, containing 50 mM $MgCl_2$). After incubation at 37° C. for signal development, the optical densities at 400 nm were determined. Competition assays were performed with the addition of increasing amounts of free NPN hapten ranging from zero up to 5 mg/well.

The ELISA assays confirmed the presence of functional antibody Fab. In a two site ELISA on NPN antigen coated plates when probed with anti-mouse kappa chain enzyme conjugate, phage supernatant generated from helper phage infection of cells carrying the pCBAK8-2b construct exhibited expected titration curves with serial two fold dilutions of phage containing antibody. The results of the two-site ELISA are shown in FIG. 10. For a signal to be generated in this assay, the phage particle must (i) have functionally associated Fd and kappa chains and (ii) be multivalent. Specificity of the particle was assessed by inhibiting binding to the plate in the presence of increasing concentrations free hapten. The generated phage particles exhibited binding to solid phase of the ELISA and could be inhibited by addition of hapten as shown in FIG. 11. Complete inhibition was achieved when 5 ng of free NPN hapten was used in the assay. Helper phage did not give a signal in the ELISA.

c. Antigen Specific Precipitation of Phage

Phage supernatant from XL1-Blue was transformed with the pCBAK8-2b dicistronic expression vector prepared in Example 3 (1 ml) was incubated with BSA-NPN conjugate (10 μl, 2 mg/ml) for 18 hours at 4° C. The mixture was then pelleted by centrifugation at 3000 rpm on a bench top centrifuge and the appearance of precipitate noted. Helper phage was used as a control. The pellet was washed repeatedly in cold PBS (5×3 ml/wash) and then resuspended in LB (0.5 ml). The solubilized precipitates were added to fresh XLI-Blue cells (0.5 ml of overnight culture), incubated for 1 hour at 37° C. and aliquots plated out on LB agar containing chloramphenicol (30 μg/ml). Colonies were selected randomly. Colony lifts on nitrocellulose were treated with lysozyme to digest the cell wall, briefly treated with chloroform to breakdown the outer membrane, blocked in BSA 1% in PBS and incubated with ^{125}I labelled BSA-NPN antigen. After several washes in PBS (containing 0.05% Tween-20), film was exposed to the washed and dried filter overnight at −70° C. and the autoradiographs were then developed.

Precipitates were obtained with antibody containing phage but not helper phage in the presence of BSA-NPN. In addition, the particles retained infectivity on subsequent incubation with bacterial cells carrying the F' episome and generated 4×10^5 colonies from a single solubilized precipitate.

Additionally, DNA restriction analysis was carried out to determine the presence of heavy and light chain inserts. DNA restriction analysis of the clones revealed the presence of a. Xho and Xba I fragment of 1.4 kb as expected for Fd-cpVIII fusion construct and kappa chain insert.

These results give additional evidence for antigen specificity and multivalency. In addition to providing immunological parameters, this precipitation offers possibilities for facile enrichment of antigen specific phage particles. In principle, phage containing specific antibodies can be highly enriched by precipitation with antigens (which may be cell surface markers, viral, bacterial as well as synthetic molecules). The washed antigen-antibody precipitates can be solubilized by the addition of excess antigen and viable

phage recovered. For the recovery of rare species an immobilized antigen may be used which opens the way for differential affinity elution.

In order to demonstrate the utility of immobilized antigen for the enrichment of clones of defined binding specificity, a panning experiment was performed. An ampicillin resistant phagemid expressing an anti-tetanus Fab as a cpVIII fusion was constructed. Rescue of this clone with helper phage produced phage encoding the ampicillin resistant phagemid which displayed the anti-tetanus Fab on their coat. These phage encoding tetanus specificity were admixed with NPN hapten encoding phage (1:100) and allowed to bind to a microtitration plate coated with tetanus toxoid. Following a one hour maintenance period, the plate was washed extensively and phage were then eluted with a low pH buffer. Infection of XL1-Blue cells in log phase growth and subsequent plating of aliquots on ampicillin and chloramphenicol allowed for direct quantitation of enrichment. Examination of over 1,000 colonies showed that ampicillin resistant colonies derived from the eluted phage exceeded chloramphenicol resistant colonies by 27 to 1. Therefore, panning enriched the phage displaying the anti-tetanus Fab by 2700 fold. This result suggests that a clone of defined specificity present at one part per million will dominate over nonspecific clones following two rounds of panning.

5. Advantages of Assembling Combinatorial Antibody Fab Libraries Along Phage Surfaces

A powerful technique for generating and selecting combinatorial Fabs, with 10^{8-9} members, is presented. In the vector described herein, the restriction cloning sites for inserting PCR generated antibody fragments have been retained as previously reported for the lambda vector. The rescue of the genes encoding the antibody Fd and kappa chains is mediated through the utilization of the f1 origin of replication leading to the synthesis and packaging of the positive strand of the vector on co-infection with helper phage. Since the 'mature' virus particle assembles by incorporating the major coat protein around the single stranded DNA as it passes through the inner membrane into the periplasmic space, not only does it capture the genetic information carried on the phagemid vector but also incorporates several copies of functional Fab along the length of the particle. On subsequent infection of hosts cells carrying the F' episome the phagemid confers resistance allowing selection of colonies on the appropriate antibiotic. In essence, the antigen recognition unit has been linked to instructions for its production.

The full power of the earlier combinatorial system could not be fully utilized since screening allowed ready access to only about 0.1–1% of the members. In the phagemid/M13 system similar size libraries are generated and all the members are accessed via affinity selection. Furthermore, unlike the lambda vector which generated monovalent Fabs, this system generates multivalent particles, thus allowing the capture of a wider range of affinities.

The unique phagemid restriction sites permit the recombination of Fd and kappa chains allowing chain replacement or shuffling. The rescue of filamentous single stranded DNA allows rapid sequencing and analysis of the genetic make up of the clone of interest. Indeed it can be envisaged that phage encoding antibody specificity may be enriched by antigen selection prior to DNA sequencing or mutagenesis. The option to further develop an iterative process of mutation followed by selection may allow the rapid generation of high affinity antibodies from germ line sequences. The process may be automated. Setting aside the potential of the system to mimic nature, the phagemid/M13 system would allow a more complete dissection of the antibody response in humans which may yield useful therapeutic and diagnostic reagents.

The membrane anchoring of the heavy chain and the compartmentalization of the kappa chain in the periplasm is the key to expressing this functional dimeric protein. The potential of this system is by no means limited to antibodies and may be extended to any protein recognition system or combination of systems containing multiple members. For example coupling of ligand and effector systems in a high avidity matrix is now possible. In a similar vein a library of ligands can be sorted against a library of receptors.

6. Randomized Mutagenesis of the CDR3 Region of a Heavy Chain Encoding Tetanus Toxoid

a. PCR Mutagenesis with Degenerate Oligonucleotides

To obtain a mutagenized heterodimer of this invention of altered specificity that would no longer recognize TT but would recognize and specifically bind to a new antigen, a method was developed to randomize only the CDR3 region of a heavy chain fragment encoded by a known nucleotide sequence. This approach is schematically diagrammed in FIG. 12 where a representative heavy chain fragment within a phagemid clone, consisting of alternating framework regions (1 through 4) shown by white blocks and complementarity determining regions (CDR) (1 through 3) shown by cross-hatched blocks and the first constant region (CH1), is subjected to two separate rounds of PCR. In the first PCR amplification reaction, the 5' end of the heavy chain beginning at framework 1 and extending to the 3' end of framework 3 is amplified. In the second PCR amplification reaction, the CDR3 region is randomly mutagenized shown by the black box. This is accomplished through the use of a pool of oligonucleotide primers synthesized with a degenerate region sandwiched between and contiguous with conserved framework 3 and 4 region sequences. The resulting amplification products, each having a randomized CDR3 region, begin at the 3' end of framework 3 and extend to the 3' end of the CH1 region. The pool of degenerate oligonucleotide primers have been designed to result in the amplification of products having a 5' end that is complementary to and will overlap with the 3' end of the products of the first PCR reaction product. Thus, the two separate PCR reaction products are pooled and subjected to a third PCR reaction in which the overlapping region between the two products is extended to result in heavy chain having a randomized CDR3 region.

A heavy chain DNA template for use in this invention was available in a clone (a phagemid vector containing heavy and light chain fragments) from a human combinatorial anti-tetanus toxoid (TT) Fab library. This library was constructed in the pCBAK-3 dicistronic expression vector for the expression of a heavy chain-cpIII fusion protein (Fd-cpIII) and a soluble light chain as described for anti-NPN in Example 2k and by Persson et al., *Proc. Natl. Acad. Sci., USA*, 88:2432–2436 (1992) and Barbas et al., *Proc. Natl. Acad. Sci., USA*, 88:7978–7982 (1992). A clone, hereinafter referred to as pCE-TT7E, was expressed as described for anti-NPN heterodimers on phage surfaces in Example 3 and subsequently screened by panning on TT-coated plates as described for anti-NPN in Example 4c . Clone pCE-TT7E exhibited a K_d towards TT on the order of 10^{-7}M and was enriched over nonspecific phage by 10^3-fold as described by Barbas et al., supra. Clone pCE-TT7E, having both heavy and light chain sequences, was used as the template DNA for the randomized mutagenesis of the CDR3 region of the

heavy chain to alter antigen binding specificity as described herein. The sequence of the heavy chain was determined as described in Example 1a(ii). Two separate PCR reactions were performed as illustrated in FIG. 12.

The first PCR reaction resulted in the amplification of the region of the heavy chain fragment in the pC3-TT7E clone beginning of framework region 1 and extending to the end of framework region 3 which is located 5' to CDR3 which is approximately 400 base pairs in length. To amplify this region, the following primer pairs were used. The 5' anti-sense oligonucleotide primer, FT3X, having the nucleotide sequence 5'-G-CAA-TAA-ACC-CTC-ACT-AAA-GGG-3' (SEQ ID NO 118), hybridized to the non-coding strand of the heavy chain corresponding to the region 5' of and including the beginning of framework 1. The 3' sense oligonucleotide primer, B7EFR3, having the nucleotide sequence 5'-TCT-CGC-ACA-ATA-ATA-CAC-GGC-3' (SEQ ID NO 119), hybridized to the coding strand of the heavy chain corresponding to the 3' end of the framework-3 region. The oligonucleotide primers were synthesized by Research Genetics (Hunstville, Ala.). The PCR reaction was performed in a 100 ul reaction containing one ug of each of oligonucleotide primers FTX3 and B7EFR3, 8 ul 2.5 mM dNTP's (dATP, dCTP, dGTP, dTTP), 1 ul Taq polymerase, 10 ng of template pCE-TT7E, and 10 ul of 10×PCR buffer purchased commercially (Promega Biotech). Two drops of mineral oil were placed on top of the admixture and 35 rounds of PCR amplification in a thermocycler were performed. The amplification cycle consisted of denaturing at 94 C. for one minute, annealing at 50° C. for one minute, followed by extension at 72° C. for two minutes. The resultant PCR amplification products were then gel purified as described in Example 1d and used in an overlap extension PCR reaction with the products of the second PCR reaction, both as described below, to recombine the two products into reconstructed heavy chains containing mutagenized CDR3 regions as illustrated in FIG. 12.

The second PCR reaction resulted in the amplification of the heavy chain from the 3' end of framework region 3 extending to the end of CH1 region which is approximately 390 base pairs in length. To amplify this region, the following primer pairs were used. The 5' anti-sense oligonucleotide primer pool, designated 7ECDR3, had the nucleotide sequence represented by the formula,

5'-GTG-TAT-TAT-TGT-GCG-AGA-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-NNS-TGG-GGC-CAA-GGG-A CC-ACG-3'

where N can be A, C, G, or T and S is either C or G (SEQ ID NO 120), wherein the 5' end of the primer pool is complementary to the 3' end of framework 3 represented by the complementary nucleotide sequence of the oligonucleotide primer B73FR3 and the 3' end of the primer pool is complementary to the 5' end of framework 4. The region between the two specified ends of the primer pool is represented by a 48-mer NNS degeneracy which ultimately encodes a diverse population of mutagenized CDR3 regions of 16 amino acid residues in length. The 3' sense oligonucleotide primer, CG1Z, as described by Persson et al., supra, having the nucleotide sequence 5'-GCATGTACTAGTTTTGTCACAAGATTTGGG-3' (SEQ ID NO 121), hybridized to the coding strand of the heavy chain corresponding to the 3' end of the CH1. The second PCR reaction was performed on the pC3-TT7E in a 100 ul reaction as described above containing one ug of each of oligonucleotide primers 7ECDR3 and CG1Z. The resultant PCR amplification product was then gel purified as described above.

One hundred nanograms of gel purified products from the first and second PCR reactions were then admixed with 1 ug each of FTX3 and CG1Z oligonucleotide primers as a primer pair in a final PCR reaction to form a complete heavy chain fragment by overlap extension as illustrated in FIG. 12. The PCR reaction admixture also contained 10 ul 10×PCR buffer, 1 ul Taq polymerase and 8 ul 2.5 mM d NTP's as described above. The PCR reaction was performed as described above. To obtain sufficient quantities of amplification product, 15 identical PCR reactions were performed. The resulting heavy chain fragments beginning at framework 1 and extending to the end of CH1 and having randomly mutagenized CDR3 regions were approximately 790 base pairs in length. The heavy chain fragment amplification products from the 15 reactions were first pooled and then gel purified as described above prior to their incorporation into a phagemid library.

b. Phagemid Library Construction

The resultant gel purified heavy chain fragments prepared in Example 6a were then digested with the restriction enzymes, Xho I and Spe I, as described in Example 2d. The resultant digested heavy chain fragments were subsequently gel purified prior to insertion into the pC3-TT7E phagemid vector clone which was previously digested with the same restriction enzymes to remove the non-mutagenized heavy chain fragment and form a linear vector. Ligation of 640 ng of the heavy chain Xho I/Spe I fragments having mutagenized CDR3 regions into two ug of the linearized pC3-TT7E phagemid vector to form circularized vectors having mutagenized CDR3 regions was performed overnight at room temperature using 10 units of BRL ligase (Gaithersburg, Md.) in BRL ligase buffer in a reaction volume of 150 uL. Five separate ligation reactions were performed to increase the size of the phage library having mutagenized CDR3 regions. Following the ligation reactions, the circularized DNA was precipitated at −20° C. for two hours by the admixture of 2 ul of 20 mg/ml glycogen, 15 ul of 3M sodium acetate at pH 5.2 and 300 ul of ethanol. DNA was then pelleted by microcentrilugation at 4° C. for 15 minutes. The DNA pellet was washed with cold 70% ethanol and dried under vacuum. The pellet was resuspended in 10 ul of water and transformed by electroporation into 300 ul of *E. coli* XL1-Blue cells as described in Example 2k to form a phage library. The total yield from the mutagenesis and transformation procedure described herein was approximately 5×10^7 transformants.

After transformation, to isolate phage on which heterodimer expression has been induced for subsequent panning on target antigens such as fluorescein, 3 ml of SOC medium (SOC was prepared by admixture of 20 g bactotryptone, 5 g yeast extract and 0.5 g NaCl in one liter of water, adjusting the pH to 7.5 and admixing 20 ml of glucose just before use to induce the expression of the Fd-cpIII and light chain heterodimer) was admixed and the culture was shaken at 220 rpm for one hour at 37 C., after which 10 ml of SB (SB was prepared by admixing 30 g tryptone, 20 g yeast extract, and 10 g Mops buffer per liter with pH adjusted to 7) containing 20 ug/ml carbenicillin and 10 ug/ml tetracycline and the admixture was shaken at 300 rpm for an additional hour. This resultant admixture was admixed to 100 ml SB containing 50 ug/ml carbenicillin and 10 ug/ml tetracycline and shaken for one hour, after which helper phage VCSM13 (10^{12} pfu) were admixed and the admixture was shaken for an additional two hours. After this time, 70 ug/ml kanamycin was admixed and maintained at 30° C. overnight. The lower temperature resulted in better heterodimer incorporation on the surface of the phage. The

supernatant was cleared by centrifugation (4000 rpm for 15 minutes in a JA10 rotor at 4° C.). Phage were precipitated by admixture of 4% (w/v) polyethylene glycol 8000 and 3% (w/v) NaCl and maintained on ice for 30 minutes, followed by centrifugation (9000 rpm for 20 minutes in a JA10 rotor at 4° C.). Phage pellets were resuspended in 2 ml of PBS and microcentrifuged for three minutes to pellet debris, transferred to fresh tubes and stored at −20° C. for subsequent screening as described below.

For determining the titering colony forming units (cfu), phage (packaged phagemid) were diluted in SB and 1 ul was used to infect 50 ul of fresh (AOD600=1) *E. coli* XLI-Blue cells grown in SB containing 10 ug/ml tetracycline. Phage and cells were maintained at room temperature for 15 minutes and then directly plated on LB/carbenicillin plates.

c. Selection of Anti-Fluorescein Heterodimers on Phage Surfaces

1) Multiple Pannings of the Phage Library Having Mutagenized CDR3 Regions

The phage library produced in Example 6b having heavy chain fragments with mutagenized CDR3 regions was panned as described herein on a microtiter plate coated with a 50 ug/ml fluorescein-BSA conjugate to screen for anti-fluorescein heterodimers. Fluorescein was conjugated to BSA according to the methods described in "Antibodies: A Laboratory Manual", eds Harlow et al., Cold Spring Harbor Laboratory, 1988.

The panning procedure described was a modification of that originally described by Parmley and Smith (Parmley et al., *Gene*, 73:30–5–318). Two to four wells of a microtiter plate (Costar 3690) were coated overnight at 4° C. with 25 µl of 50 µg/ml antigen prepared above in 0.1M bicarbonate, pH 8.6. The wells were washed twice with water and blocked by completely filling the well with 3% (w/v) bovine serum albumin (BSA) in PBS and incubating the plate at 37° C. for 1 hour. Blocking solution was shaken out, 50 µl of the phage library prepared above (typically 10^{11} cfu) was added to each well, and the plate was incubated for 2 hours at 37° C.

Phage were removed and the plate was washed once with water. Each well was then washed 10 times with TBS/Tween (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% Tween 20) over a period of 1 hour at room temperature —pipetted up and down to wash the well, each time allowing the well to remain completely filled with TBS/Tween between washings. The plate was washed once more with distilled water and adherent phage were eluted by the addition of 50 µl of elution buffer (0.1M HCl adjusted to pH 2.2 with solid glycine, containing 1 mg/ml BSA) to each well and incubation at room temperature for 10 minutes. The elution buffer was pipetted up and down several times, removed, and neutralized with 3 µl of 2M Tris base per 50 µl of elution buffer used. Eluted phage were used to infect 2 ml of fresh (OD_{600}=1) *E. coli* XLI-Blue cells for 15 minutes at room temperature, after which 10 ml of SB containing 20 µg/ml carbenicillin and 10 µg/ml tetracycline was admixed. [Aliquot (20, 10, and 1/10 µl) were removed for plating to determine the number of phage (packaged phagemids) that were eluted from the plate.] The culture was shaken for 1 hour at 37° C., after which it was added to 100 ml of SB containing 50 µg/ml carbenicillin and 10 µg/ml tetracycline and shaken for 1 hour. Then helper phage VCSM13 (10^{12} pfu) were added and the culture was shaken for an additional 2 hours. After this time, 70 µg/ml kanamycin was added and the culture was incubated at 37° C. overnight. Phage preparation and further panning were repeated as described above.

Following each round of panning, the percentage yield of phage must be determined, where % yield—(number of phage eluted/number of phage applied) ×100.

As an alternative to elution with acid, phage bound to the wells of the microtiter plate were eluted by admixing 50 ul of a solution of 10^{-5}M fluorescein diluted in PBS followed by a maintenance period of one hour at 37 C. The solution was then pipetted up and down to wash the wells. The resultant eluate was transferred to 2 ml of fresh *E. coli* XLI-Blue cells for infection as described above for preparing phage and further panning. In subsequent rounds of panning, phage were eluted with 10^{-6}M fluorescein.

The results of the amount of phage that were specifically bound to fluorescein-coated wells over four consecutive rounds of panning and eluted with acid or with fluorescein alone are shown below in Table 8. Comparable yields of phage on which heterodimers were expressed that bound specifically to fluorescein were achieved with either elution protocol. Approximately 20 clones of the 5×10^7 clones resulting from the mutagenesis and transformation exhibited specificity of binding towards fluorescein-coated wells. These data confirm that mutagenesis of the CDR3 region as described in this invention resulted in the altering of a heterodimer which initially specifically bound to TT to one that specifically bound fluorescein.

TABLE 8

	Phage Eluted	
	Acid Elution	Fluorescein Elution
round 1	5.6×10^5/well	4.7×10^5/well
round 2	4.6×10^6/well	5.6×10^5/well
round 3	3.75×10^5/well	1.35×10^6/well
round 4	1.31×10^6/well	4.0×10^6/well

2) Preparation of Soluble Heterodimers for Characterizing Binding Specificity to Fluorescein

In order to further characterize the specificity of the mutagenized heterodimers expressed on the surface of phage as described above, soluble heterodimers were prepared and analyzed in ELISA assays on fluorescein-coated plates, by competive ELISA with increasing concentrations of soluble fluorescein-BSA and also by fluorescence quenching assays. The latter assays were performed as described in "Fluorescein Hapten: An Immunological Probe", ed E. W. Voss, CRC Press, Inc. pp 52–54, 984.

To prepare soluble heterodimers, phagemid DNA from positive clones was isolated and digested with SpeI and NheI. Digestion with these enzymes produces compatible cohesive ends. The 4.7-kb DNA fragment lacking the gIII portion was gel-purified (0.6% agarose) and self-ligated. Transformation of *E. coli* XLI-Blue afforded the isolation of recombinants lacking the gIII fragment. Clones were examined for removal of the gIII fragment by XhoI/XbaI digestion, which should yield an 1.6-kb fragment. Clones were grown in 100 ml SB containing 50 µg/ml carbenicillin and 20 mM $MgCl_2$ at 37° C. until an OD_{600} of 0.2 was achieved. IPTG (1 mM) was added and the culture grown overnight at 30° C. (growth at 37° C. provides only a light reduction in heterodimer yield). Cells were pelleted by centrifugation at 4000 rpm for 15 minutes in a JA10 rotor at 4° C. Cells were resuspended in 4 ml PBS containing 34 µg/ml phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication on ice (2–4 minutes at 50% duty). Debris was pelleted by centrifugation at 14,000 rpm in a JA20 rotor at

4° C. for 15 minutes. The supernatant was used directly for ELISA analysis and was stored at −20° C. For the study of a large number of clones, 10-ml cultures provided plenty of heterodimer for analysis. In this case, sonications were performed in 2 ml of buffer.

The soluble heterodimers prepared above were assayed by ELISA. For this assay, 1 ug/well of fluorescein-BSA solution was admixed to individual wells of a microtiter plate and maintained at 4° C. overnight to allow the protein solution to adhere to the walls of the well. After the maintenance period, the wells were washed one time with PBS and thereafter maintained with a solution of 3% BSA to block nonspecific sites on the wells. The plates were maintained at 37° C. for one hour after which time the plates were inverted and shaken to remove the BSA solution. Soluble heterodimers prepared above were then admixed to each well and maintained at 37° C. for one hour to form an immunoreaction products. Following the maintenance period, the wells were washed 10 times with PBS to remove unbound soluble antibody and then maintained with a secondary goat anti-human FAB conjugated to alkaline phosphatase diluted in PBS containing 1% BSA. The wells were maintained at 37° C. for one hour after which the wells were was 10 times with PBS followed by development with p-nitrophenyl phosphate.

Immunoreactive heterodimers as determined in the above ELISA were then analyzed by competition ELISA to determine the affinity of the mutagenized heterodimers. The ELISA was performed as described above with increasing concentrations of soluble fluorescein-BSA ranging in concentration from 10^{-9}M up to 10^{-5}M in concentration admixed in the presence of the soluble heterodimers. Maximal inhibition of binding was achieved at a concentration of 10^{-6}M free antigen with a half-maximal inhibition obtained with approximately 10^{-7}M free antigen. Thus, the mutagenized heterodimers of this invention specifically recognize and bind to fluorescein. Additional experiments were performed to confirm that the mutagenized heterodimers no longer recognized the TT to which they nonmutagenized heterodimer originally bound. Fluorescence quenching assays were also performed to confirm the specificity of binding of the mutagenized heterodimers. Soluble heterodimers prepared from phage that were either eluted with acid or with fluorescein alone were equally effective at binding fluorescein by any of the aforementioned approaches. The invention of mutagenesis of the CDR3 region of the heavy chain of a heterodimer described herein thus resulted in the alteration of binding specificity from TT to fluorescein.

Thus, the above example illustrates a method according to the present invention for mutagenizing the complementarity determining region (CDR) of an immunoglobulin gene, and also illustrates oligonucleotides useful therefor.

In one embodiment, therefore, an oligonucleotide is contemplated that is useful as a primer in a polymerase chain reaction (PCR) for inducing mutagenesis in a complementarity determining region (CDR) of an immunoglobulin gene. The oligonucleotide has 3' and 5' termini and comprises (1) a nucleotide sequence at its 3' terminus capable of hybridizing to a first framework region of an immunoglobulin gene, (2) a nucleotide sequence at its 5' terminus capable of hybridizing to a second framework region of an immunoglobulin gene, and (3) a nucleotide sequence between the 3' and 5' termini adapted for introducing mutations during a PCR into the CDR region between the first and second framework regions of the immunoglobulin gene, thereby mutagenizing the CDR region.

Insofar as immunoglobulin genes have three CDR regions on both the heavy chain and the light chain of an immunoglobulin, each separated by a distinctive framework region, it is to be understood that the above example is readily applicable to introducing mutations into a specific CDR by selection of the above 5' and 3' nucleotide sequences as to hybridize to the framework regions flanking the targeted CDR. Thus the above first and second framework sequences can be the conserved sequences flanking CDR1, CDR2 or CDR3 on either the heavy or light chain. Exemplary and preferred is the CDR3 of the human immunoglobulin heavy chain.

The length of the 3' and 5' terminal nucleotide sequences of a subject mutagenizing oligonucleotide can vary in length as is well known, so long as the length provides a stretch of nucleotides complementary to the target framework sequences as to hybridize thereto. In the case of the 3' terminal nucleotide sequence, it must be of sufficient length and complementarity to the target framework region located 3' to the CDR region to be mutagenized as to hybridize and provide a 3' hydroxyl terminus for initiating a primer extension reaction. In the case of the 5' terminal nucleotide sequence, it must be of sufficient length and complementarity to the target framework region located 5' to the CDR region to be mutagenized as to provide a means for hybridizing in a PCR overlap extension reaction as described above to assemble the complete immunoglobulin heavy or light chain.

Framework regions flanking a CDR are well characterized in the immunological arts, and include known nucleotide sequences or consensus sequences as described elsewhere herein. Where a single, preselected immunoglobulin gene is to be mutagenized, the framework-defined sequences flanking a particular CDR are known, or can be readily determined by nucleotide sequencing protocols. Where a repertoire of immunoglobulin genes are to be mutagenized, the framework-derived sequences are preferably conserved, as described elsewhere herein.

Preferably, the length of the 3' and 5' terminal nucleotide sequences are each at least 6 nucleotides in length, and can be up to 50 or more nucleotides in length, although these lengths are unnecessary to assure accurate and reproducible hybridization. Preferred are lengths in the range of 12 to 30 nucleotides, and typically are about 18 nucleotides.

A particularly preferred framework-defined nucleotide sequence for use as a 3' terminus nucleotide sequence has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122).

A particularly preferred framework-defined nucleotide sequence for use as a 5' terminus nucleotide sequence has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123).

The nucleotide sequence located between the 3' and 5' termini adapted for mutagenizing a CDR can be any nucleotide sequence, insofar as the novel sequence will be incorporated by the above methods. However, the present approach provides a means to produce a large population of mutagenized CDR's in a single PCR reaction by the use of a population of redundant sequences defining randomized or nearly randomized nucleotides in the CDR region to be mutagenized.

A preferred oligonucleotide comprises a nucleotide sequence between the above described 3' and 5' termini that is represented by the formula: $[NNS]_n$ or $[NNK]_n$, wherein N can independently be any nucleotide, where S is G or C, K is G or T, and where n is from 3 to about 24. In preferred embodiments the preferred oligonucleotides have the formula:

5'-GTGTATTATTGTGCGAGA[NNS]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO: 124) and 5'-GTGTATTATTGTGC-GAGA[NNK]$_n$TGGGGCCAAGGGACCACG-3' (SEQ ID NO: 125).

Exemplary and particularly preferred is the oligonucleotide where with the formula $[NNS]_n$ and n is 16, such that the oligonucleotide represents a large population of redundant oligonucleotide sequences.

The invention also contemplates a mutagenesis method for altering the immunological specificity of a cloned immunologlobulin gene. The method provides direct mutagenesis in a preselected CDR of an immunoglobulin gene which comprises subjecting a recombinant DNA molecule (rDNA) containing the cloned immunoglobulin gene having a target CDR to PCR conditions suitable for amplifying a preselected region of the CDR. In the method, the rDNA molecule is subjected to PCR conditions that include a PCR primer oligonucleotide as described above constituting the first primer in a PCR primer pair as is well known to produce an amplified PCR product that is derived from the preselected CDR region but that includes the nucleotide sequences of the PCR primer. The second oligonucleotide in the PCR amplifying conditions can be any PCR primer derived from the immunoglobulin gene to be mutagenized, as described herein.

Preferred are methods using an oligonucleotide of this invention as described above.

The foregoing is intended as illustrative of the present invention but not limiting. Numerous variations and modifications can be effected without departing from the true spirit and scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 125

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 173 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCCGCAAAT TCTATTTCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC TACGGCAGCC 60

GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCC AGGTGAAACT GCTCGAGATT 120

TCTAGACTAG TTACCCGTAC GACGTTCCGG ACTACGGTTC TTAATAGAAT TCG 173

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 173 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCGACGAATT CTATTAAGAA CCGTAGTCCG GAACGTCGTA CGGGTAACTA GTCTAGAAAT 60

CTCGAGCAGT TTCACCTGGG CCATGGCTGG TTGGGCAGCG AGTAATAACA ATCCAGCGGC 120

TGCCGTAGGC AATAGGTATT TCATTATGAC TGTCTCCTTG AAATAGAATT TGC 173

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGAATTCTAA ACTAGTCGCC AAGGAGACAG TCATAATGAA ATACCTATTG CCTACGGCAG 60

CCGCTGGATT GTTATTACTC GCTGCCCAAC CAGCCATGGC CGAGCTCGTC AGTTCTAGAG 120

TTAAGCGGCC G 131

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TCGACGGCCG CTTAACTCTA GAACTGACGA GCTCGGCCAT GGCTGGTTGG GCAGCGAGTA 60

ATAACAATCC AGCGGCTGCC GTAGGCAATA GGTATTTCAT TATGACTGTC TCCTTGGCGA 120

CTAGTTTAGA ATTCAAGCT 139

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

Ala Gln Pro Ala Met
20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Erwinia carotovora

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Leu Ala
1 5 10 15

-continued

Ala Gln Pro Ala Gln Pro Ala Met Ala
20 25

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(v i) ORIGINAL SOURCE:
(A) ORGANISM: Erwinia carotovora

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Lys Ser Leu Ile Thr Pro Ile Ala Ala Gly Leu Leu Leu Ala Phe
1 5 10 15

Ser Gln Tyr Ser Leu Ala
20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Lys Ile Lys Thr Gly Ala Arg Ile Leu Ala Leu Ser Ala Leu Thr
1 5 10 15

Thr Met Met Phe Ser Ala Ser Ala Leu Ala Lys Ile
20 25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Met Lys Arg Asn Ile Leu Ala Val Ile Val Pro Ala Leu Leu Val
1 5 10 15

Ala Gly Thr Ala Asn Ala Ala Glu
20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

-continued

```
Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr
1               5                  10                  15

Pro Val Thr Lys Ala Arg Thr
            20
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Met Ile Thr Leu Arg Lys Leu Pro Leu Ala Val Ala Val Ala Ala
1               5                  10                  15

Gly Val Met Ser Ala Gln Ala Met Ala Val Asp
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Lys Ala Thr Lys Leu Val Leu Gly Ala Val Ile Leu Gly Ser Thr
1               5                  10                  15

Leu Leu Ala Gly Cys Ser
            20
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

-continued

```
Met Lys Lys Ser Leu Val Leu Lys Ala Ser Val Ala Val Ala Thr Leu
1               5                  10                  15

Val Pro Met Leu Ser Phe Ala
            20
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: peptide

(i i i) HYPOTHETICAL: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Lys Lys Leu Leu Phe Ala Ile Pro Leu Val Val Pro Phe Tyr Ser
1               5                  10                  15

His Ser
```

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 211 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(i i) MOLECULE TYPE: protein

(i i i) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
Pro Phe Val Cys Glu Tyr Gln Gly Gln Gly Gln Ser Ser Asp Leu Pro
1               5                  10                  15

Gln Pro Pro Val Asn Ala Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly
            20                  25                  30

Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly
        35                  40                  45

Ser Glu Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp
    50                  55                  60

Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala
65                  70                  75                  80

Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val
                85                  90                  95

Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser
            100                 105                 110

Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn
        115                 120                 125

Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn
    130                 135                 140

Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg
145                 150                 155                 160

Pro Phe Val Phe Ser Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys
                165                 170                 175

Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val
            180                 185                 190

Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn
        195                 200                 205
```

-continued

Lys Glu Ser
210

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ala Glu Gly Asp Asp Pro Ala Lys Ala Ala Phe Asn Ser Leu Gln Ala
1 5 10 15

Ser Ala Thr Glu Tyr Ile Gly Tyr Ala Trp Ala Met Val Val Val Ile
20 25 30

Val Gly Ala Thr Ile Gly Ile Lys Leu Phe Lys Lys Phe Thr Ser Lys
35 40 45

Ala Ser
50

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AAUCUUGGAG GCUUUUUUAU GGUUCGUUCU 30

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

UAACUAAGGA UGAAAUGCAU GUCUAAGACA 30

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

UCCUAGGAGG UUUGACCUAU GCGAGCUUUU 30

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: RNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AUGUACUAAG GAGGUUGUAU GGAACAACGC 30

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GGCCGCAAAT TCTATTTCAA GGAGACAGTC AT 32

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATT 36

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

-continued

GTTATTACTC GCTGCCCAAC CAGCCATGGC CC 32

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CAGTTTCACC TGGGCCATGG CTGGTTGGG 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CAGCGAGTAA TAACAATCCA GCGGCTGCCG TAGGCAATAG 40

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GTATTTCATT ATGACTGTCT CCTTGAAATA GAATTTGC 38

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AGGTGAAACT GCTCGAGATT TCTAGACTAG TTACCCGTAC 40

(2) INFORMATION FOR SEQ ID NO:29:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGGAACGTCG TACGGGTAAC TAGTCTAGAA ATCTCGAG 38

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GACGTTCCGG ACTACGGTTC TTAATAGAAT TCG 33

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TCGACGAATT CTATTAAGAA CCGTAGTC 28

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TGAATTCTAA ACTAGTCGCC AAGGAGACAG TCAT 34

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid

-continued

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATT 36

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTTATTACTC GCTGCCCAAC CAGCCATGGC C 31

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GAGCTCGTCA GTTCTAGAGT TAAGCGGCCG 30

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GTATTTCATT ATGACTGTCT CCTTGGCGAC TAGTTTAGAA TTCAAGCT 48

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

-continued

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CAGCGAGTAA TAACAATCCA GCGGCTGCCG TAGGCAATAG 40

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TGACGAGCTC GGCCATGGCT GGTTGGG 27

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TCGACGGCCG CTTAACTCTA GAAC 24

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:40:

AGGTSMARCT KCTCGAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AGGTCCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AGGTCCAGCT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AGGTCCAGCT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:44:

AGGTCCAGCT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:45:

AGGTCCAACT GCTCGAGTCT GG 22

-continued

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AGGTCCAACT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:47:

AGGTCCAACT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AGGTCCAACT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGGTNNANCT NCTCGAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO:50:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GCCCAAGGAT GTGCTCACC 19

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CTATTAGAAT TCAACGGTAA CAGTGGTGCC TTGGCCCCA 39

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CTATTAACTA GTAACGGTAA CAGTGGTGCC TTGCCCCA 38

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CTCAGTATGG TGGTTGTGC 19

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

-continued

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GCTACTAGTT TTGATTTCCA CCTTGG 26

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CAGCCATGGC CGACATCCAG ATG 23

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:56:

AATTTTACTA GTCACCTTGG TGCTGCTGGC 30

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:57:

TATGCAACTA GTACAACCAC AATCCCTGGG CACAATTTT 39

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

-continued

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:58:

AGGCTTACTA GTACAATCCC TGGCACAAT 30

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:59:

CCAGTTCCGA GCTCGTTGTG ACTCAGGAAT CT 32

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CCAGTTCCGA GCTCGTGTTG ACGCAGCCGC CC 32

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:61:

CCAGTTCCGA GCTCGTGCTC ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

-continued

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CCAGTTCCGA GCTCCAGATG ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:63:

CCAGATGTGA GCTCGTGATG ACCCAGACTC CA 32

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:64:

CCAGATGTGA GCTCGTCATG ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:65:

CCAGATGTGA GCTCTTGATG ACCCAAACTC AA 32

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:66:

CCAGATGTGA GCTCGTGATA ACCCAGGATG AA 32

-continued

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC 32

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CCGCCGTCTA GAACACTCAT TCCTGTTGAA GCT 33

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:69:

CCGCCGTCTA GAACATTCTG CAGGAGACAG ACT 33

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:70:

CCAGTTCCGA GCTCGTGATG ACACAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

-continued

(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GCGCCGTCTA GAATTAACAC TCATTCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CTATTAACTA GTAACGGTAA CAGTGGTGCC TTGCCCCA 38

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:73:

AGGCTTACTA GTACAATCCC TGGGCACAAT 30

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:74:

GCCGCTCTAG AACACTCATT CCTGTTGAA 29

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:75:

AGGTNNANCT NCTCGAGTCT GC 22

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:76:

AGGTNNANCT NCTCGAGTCA GC 22

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:77:

GTGCCAGATG TGAGCTCGTG ATGACCCAGT CTCCA 35

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:78:

TCCTTCTAGA TTACTAACAC TCTCCCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GCATTCTAGA CTATTATGAA CATTCTGTAG GGGC 34

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:80:

CTGCACAGGG TCCTGGGCCG AGCTCGTGGT GACTCAG 37

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:81:

AGNTGCANNT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GTGGGCATGT GTGAGTTGTG TCACTAGTTG GGGTTTTGAG CTC 43

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:83:

-continued

AGCATCACTA GTACAAGATT TGGGCTC 27

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:84:

AGGTGCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:85:

AGGTGCAGCT GCTCGAGTCG GG 22

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:86:

AGGTGCAACT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:87:

AGGTGCAACT GCTCGAGTCG GG 22

-continued

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:88:

TCCTTCTAGA TTACTAACAC TCTCCCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:89:

CTGCACAGGG TCCTGGGCCG AGCTCGTGGT GACTCAG 37

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:90:

GCATTCTAGA CTATTAACAT TCTGTAGGGG C 31

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:91:

ACCCAAGGAC ACCCTCATG 19

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs

-continued

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CTCAGTATGG TGGTTGTGC 19

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GTCTCACTAG TCTCCACCAA GGGCCCATCG GTC 33

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:94:

ATATACTAGT GAGACAGTGA CCAGGGTTCC TTGGCCCCA 39

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:95:

ACGTCTAGAT TCCACCTTGG TCCC 24

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GCATACTAGT CTATTAACAT TCTGTAGGGG C 31

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:97:

CCGGAATTCT TATCATTTAC CCGGAGA 27

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:98:

TCTGCACTAG TTGGAATGGG CACATGCAG 29

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 798 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GGCCGCAAAT TCTATTTCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC TACGGCAGCC 60

GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCC AGGTGAAACT GCTCGAGTCA 120

GGACCTGGCC TCGTGAAACC TTCTCAGTCT CTGTCTCTCA CCTGCTCTGT CACTGACTAC 180

TCCATCACCA GTGCTTATTA CTGGAACTGG ATCCGGCAGT TTCCAGGAAA CAAACTGGAA 240

TGGATGGGCT ACATAAGCTA CGACGGTGTC AATAAGTATG ATCCATCTCT CAAGAATCGA 300

ATCTCCATCA CTCGTGACAC ATCTAACAAT CAGTTTTTCC AGAAGTTGAT TTCTGTGACT 360

TCTGAGGACA CAGGAACATA TGACTGTTCA AGAGGGACTA GGGCCTCTGC TATGGACTAC 420

TGGGGTCAAG GAATTTCAGT CACCGTCTCC TCAGCCAAAA CGACACCCCC ATCTGTCTAT 480

-continued

```
CCACTGGCCC CTGGATCTGC TGCCCAAACT AACTCCATGG TGACCCTGGG ATGCCTGGTC    540
AAGGGCTATT TCCCTGAGCC AGTGACAGTG ACCTGGAACT CTGGATCCCT GTCCAGCGGT    600
GTGCACACCT TCCCAGCTGT CCTGCAGTCT GACCTCTACA CTCTGAGCAG CTCAGTGACT    660
GTCCCCTCCA GCCCTCGGCC CAGCGAGACC GTCACCTGCA ACGTTGCCCA CCCGGCCAGC    720
AGCACCAAGG TGGACAAGAA AATTGTGCCC AGGGATTGTA CTAGTTACCC GTACGACGTT    780
CCGGACTACG GTTCTTAA                                                  798
```

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 194 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:100:

```
TGAATTCTAA ACTAGTCGCC AAGGAGACAG TCATAATGAA ATACCTATTG CCTACGGCAG     60
CCGCTGGATT GTTACTCGCT GCCCAACCAG CCATGGCCGA GCTCCAGATG ACCCAGTCTC    120
CAGCCTCCCT ATCTGCATCT GTGGGAGAAA CTGTCACCAT CACATGTCGA TCAAGTGAGA    180
ATATTACAAT TACT                                                      194
```

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 333 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:101:

```
CTGATGCTGC ACCAACTGTA TCCATCTTCC CACCATCCAG TGAGCAGTTA ACATCTGGAG     60
GTGCCTCAGT CGTGTGCTTC TTGAACAACT TCTACCCCAA AGACTACAAT GTCAAGGGGA    120
AGATTGATGG CAGTGAACGA CAAAATGGCG TCCTGAACAG TTGGACTGAT CAGGACAGCA    180
AAGACAGCAC CTACAGCATG AGCAGCACCC TCACGTTGAC CAAGGACGAG TATGAACGAC    240
ATAACAGCTA TACCTGTGAT GCCACTCACA AGACATCAAC TTCACCCATT GTCAAGAGCT    300
TCAACAGGAA TGAGTGTTAA TTCTAGACGG CGC                                  333
```

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:102:

GCTGAGGGTG ACGATCCCGC AAAAGCGGCC TTTAACTCCC TGCAAGCCTC AGCGACCGAA 60

TATATCGGTT ATGCGTGGGC GATGGTTGTT GTCATTGTCG GCGCAACTAT CGGTATCAAG 120

CTGTTTAAGA AATTCACCTC GAAAGCAAGC 150

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GTGCCCAGGG ATTGTACTAG TGCTGAGGGT GACGAT 36

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:104:

ACTCGAATTC TATCAGCTTG CTTTCGAGGT GAA 33

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:105:

AGGTCCAGCT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GTCACCCTCA GCACTAGTAC AATCCCTGGG CAC 33

(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GAGACGACTA GTGGTGGCGG TGGCTCTCCA TTCGTTTGTG AATATCAA 48

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:108:

TTACTAGCTA GCATAATAAC GGAATACCCA AAAGAACTGG 40

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:109:

TATGCTAGCT AGTAACACGA CAGGTTTCCC GACTGG 36

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:110:

-continued

ACCGAGCTCG AATTCGTAAT CATGGTC 27

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 186 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GTGCCCAGGG ATTGTACTAG TGCTGAGGGT GACGATCCCG CAAAAGCGGC CTTTAACTCC 60

CTGCAAGCCT CAGCGACCGA ATATATCGGT TATGCGTGGG CGATGGTTGT TGTCATTGTC 120

GGCGCAACTA TCGGTATCAA GCTGTTTAAG AAATTCACCT CGAAAGCAAG CTGATAGAAT 180

TCGAGT 186

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 666 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:112:

CCATTCGTTT GTGAATATCA AGGCCAAGGC CAATCGTCTG ACCTGCCTCA ACCTCCTGTC 60

AATGCTGGCG GCGGCTCTGG TGGTGGTTCT GGTGGCGGCT CTGAGGGTGG TGGCTCTGAG 120

GGTGGCGGTT CTGAGGGTGG CGGCTCTGAG GGAGGCGGTT CCGGTGGTGG CTCTGGTTCC 180

GGTGATTTTG ATTATGAAAA GATGGCAAAC GCTAATAAGG GGGCTATGAC CGAAAATGCC 240

GATGAAAACG CGCTACAGTC TGACGCTAAA GGCAAACTTG ATTCTGTCGC TACTGATTAC 300

GGTGCTGCTA TCGATGGTTT CATTGGTGAC GTTTCCGGCC TTGCTAATGG TAATGGTGCT 360

ACTGGTGATT TTGCTGGCTC TAATTCCCAA ATGGCTCAAG TCGGTGACGG TGATAATTCA 420

CCTTTAATGA ATAATTTCCG TCAATATTTA CCTTCCCTCC CTCAATCGGT TGAATGTCGC 480

CCTTTTGTCT TTAGCGCTGG TAAACCATAT GAATTTTCTA TTGATTGTGA CAAAATAAAC 540

TTATTCGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 600

TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 660

TATTAT 666

(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 708 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

-continued

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:113:

GAGACGACTA GTGGTGGCGG TGGCTCTCCA TTCGTTTGTG AATATCAAGG CCAAGGCCAA 60

TCGTCTGACC TGCCTCAACC TCCTGTCAAT GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT 120

GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA 180

GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT GATTTTGATT ATGAAAAGAT GGCAAACGCT 240

AATAAGGGGG CTATGACCGA AAATGCCGAT GAAAACGCGC TACAGTCTGA CGCTAAAGGC 300

AAACTTGATT CTGTCGCTAC TGATTACGGT GCTGCTATCG ATGGTTTCAT TGGTGACGTT 360

TCCGGCCTTG CTAATGGTAA TGGTGCTACT GGTGATTTTG CTGGCTCTAA TTCCCAAATG 420

GCTCAAGTCG GTGACGGTGA TAATTCACCT TTAATGAATA ATTTCCGTCA ATATTTACCT 480

TCCCTCCCTC AATCGGTTGA ATGTCGCCCT TTTGTCTTTA GCGCTGGTAA ACCATATGAA 540

TTTTCTATTG ATTGTGACAA AATAAACTTA TTCCGTGGTG TCTTTGCGTT TCTTTTATAT 600

GTTGCCACCT TTATGTATGT ATTTTCTACG TTTGCTAACA TACTGCGTAA TAAGGAGTCT 660

TAATCATGCC AGTTCTTTTG GGTATTCCGT TATTATGCTA GCTAGTAA 708

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 201 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:114:

TATGCTAGCT AGTAACACGA CAGGTTTCCC GACTGGAAAG CGGGCAGTGA GCGCAACGCA 60

ATTAATGTGA GTTAGCTCAC TCATTAGGCA CCCCAGGCTT TACACTTTAT GCTTCCGGCT 120

CGTATGTTGT GTGGAATTGT GAGCGGATAA CAATTTCACA CAGGAAACAG CTATGACCAT 180

GATTACGAAT TCGAGCTCGG T 201

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:115:

AGGTCCAGCT TCTCGAGTCT GGACCTGGCC TCGTGAAACC TTCTCAGTCT CTGTCTCTCA 60

CCTGCTCTGT CACTGACTAC TCCATCACCA GTGCTTATTA CTGGAACTGG ATCCGGCAGT 120

TTCCAGGAAA CAAACTGGAA TGGATGGGCT ACATAAGCTA CGACGGTGTC AATAAGTATG 180

ATCCATCTCT CAAGAATCGA ATCTCCATCA CTCGTGACAC ATCTAACAAT CAGTTTTTCC 240

AGAAGTTGAT TTCTGTGACT TCTGAGGACA CAGGAACATA TGACTGTTCA AGAGGGACTA 300

-continued

```
GGGCCTCTGC TATGGACTAC TGGGGTCAAG GAATTTCAGT CACCGTCTCC TCAGCCAAAA    360
CGACACCCCC ATCTGTCTAT CCACTGGCCC CTGGATCTGC TGCCCAAACT AACTCCATGG    420
TGACCCTGGG ATGCCTGGTC AAGGGCTATT TCCCTGAGCC AGTGACAGTG ACCTGGAACT    480
CTGGATCCCT GTCCAGCGGT GTGCACACCT TCCCAGCTGT CCTGCAGTCT GACCTCTACA    540
CTCTGAGCAG CTCAGTGACT GTCCCCTCCA GCCCTCGGCC CAGCGAGACC GTCACCTGCA    600
ACGTTGCCCA CCCGGCCAGC AGCACCAAGG TGGACAAGAA AATTGTGCCC AGGGATTGTA    660
CTAGTGCTGA GGGTGACGAT CCCGCAAAAG CGGCCTTTAA CTCCCTGCAA GCCTCAGCGA    720
CCGAATATAT CGGTTATGCG TGGGCGATGG TTGTTGTCAT TGTCGGCGCA ACTATCGGTA    780
TCAAGCTGTT TAAGAAATTC ACCTCGAAAG CAAGCTGATA GAATTCGAGT               830
```

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 260 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:116:

```
ATGAAATACC TATTGCCTAC GGCAGCCGCT GGATTGTTAT TACTCGCTGC CCAACCAGCC     60
ATGGCCCAGG TGAAACTGCT CGAGATTTCT AGACTAGTGC TGAGGGTGAC GATCCCGCAA    120
AAGCGGCCTT TAACTCCCTG CAAGCCTCAG CGACCGAATA TATCGGTTAT GCGTGGGCGA    180
TGGTTGTTGT CATTGTCGGC GCAACTATCG GTATCAAGCT GTTTAAGAAA TTCACCTCGA    240
AAGCAAGCTG ATAGAATTCG                                               260
```

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:117:

```
GTACGCGCCC TGTAGCGGCG CATTAAGCGC GGCGGGTGTG GTGGTTACGC GCAGCGTGAC     60
CGCTACACTT GCCAGCGCCC TAGCGCCCGC TCCTTTTGCT TTCTTCCCTT CCTTTCTCGC    120
CACGTTCGCC GGCTTTCCCC GTCAAGCTCT AAATCGGGGG CTCCCTTTAG GGTTCCGATT    180
TAGTGCTTTA CGGCACCTCG ACCCCAAAAA ACTTGATTAG GGTGATGGTT CACGTAGTGG    240
GCCATCGCCC TGATAGACGG TTTTTCGCCC TTTGACGTTG GAGTCCACGT TCTTTAATAG    300
TGGACTCTTG TTCCAAACTG GAACAACACT CAACCCTATC TCGGTCTATT CTTTTGATTT    360
ATAAGGGATT TTGCCGATTT CGGCCTATTG GTTAAAAAAT GAGCTGATTT AACAAAAATT    420
TAACGCGAAT TTTAACAAAA TATTAACGTT TACAATTTAA A                        461
```

-continued

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:118:

GCAATAAACC CTCACTAAAG GG 22

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:119:

TCTCGCACAA TAATACACGG C 21

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 84 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:120:

GTGTATTATT GTGCGAGANN SNNSNNSNNS NNSNNSNNSN NSNNSNNSNN SNNSNNSNNS 60

NNSNNSTGGG GCCAAGGGAC CACG 84

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:121:

GCATGTACTA GTTTTGTCAC AAGATTTGGG 30

(2) INFORMATION FOR SEQ ID NO:122:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:122:

TGGGGCCAAG GGACCACG 18

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:123:

GTGTATTATT GTGCGAGA 18

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(i x) FEATURE:
(A) NAME/KEY: repeat_region
(B) LOCATION: 19..21
(D) OTHER INFORMATION: /rpt_type="tandem"
/ note="NNS can be repeated from 3 to about 24 times."

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:124:

GTGTATTATT GTGCGAGANN STGGGGCCAA GGGACCACG 39

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) HYPOTHETICAL: NO

(i v) ANTI-SENSE: NO

(i x) FEATURE:
(A) NAME/KEY: repeat_region
(B) LOCATION: 19..21
(D) OTHER INFORMATION: /rpt_type="tandem"

-continued

/ note="NNK can be repeated from 3 to about 24 times."

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:125:

GTGTATTATT GTGCGAGANN KTGGGGCCAA GGGACCACG 39

What is claimed is:

1. A method for producing a library of complementarity determining region (CDR)-mutagenized phage-displayed immunoglobulin heterodimers, the method comprising the steps of:

1) amplifying a CDR portion of a template immunoglobulin variable domain gene selected from the group consisting of a template immunoglobulin heavy chain variable domain gene and a template immunoglobulin light chain variable domain gene, wherein said template immunoglobulin heavy and light chain genes have a framework region and said CDR portion and encode respective heavy and light chain variable domain polypeptides, and wherein said amplifying is by polymerase chain reaction (PCR) using a PCR primer oligonucleotide for mutagenizing a preselected nucleotide region in said CDR portion, thereby forming a library of amplified CDR-mutagenized immunoglobulin gene fragments, said PCR primer oligonucleotide having 3' and 5' termini and comprising:

a) a nucleotide sequence at the 3' terminus capable of hybridizing to a first framework region of said selected template immunoglobulin variable domain gene;

b) a nucleotide sequence at the 5' terminus capable of hybridizing to a second framework region of said selected template immunoglobulin variable domain gene; and

c) a nucleotide sequence between the 3' and 5' termini according to the formula selected from the group consisting of:

$[NNS]_n$ and $[NNK]_n$,

wherein N is independently any nucleotide, S is G or C and K is G or T, and n is 3 to about 24, the 3' and 5' terminal nucleotide sequences having a length of about 6 to 50 nucleotides, or an oligonucleotide having a sequence complementary thereto;

2) inserting individual members of the library of amplified CDR-mutagenized immunoglobulin gene fragments formed in step (1) into a dicistronic phagemid expression vector comprising immunoglobulin heavy and light chain variable domain genes that lack the immunoglobulin gene portion corresponding to the fragment to be inserted, wherein upon insertion said vector is capable of expressing heavy and light chain variable domain polypeptides encoded by said vector, thereby forming a library of dicistronic expression vectors containing amplified CDR-mutagenized immunoglobulin gene fragments; and

3) expressing said immunoglobulin heavy and light chain genes in the library of dicistronic expression vectors formed in step (2) whereby said encoded heavy and light chain variable domain polypeptides assemble on the surface of a phage to form a phage-displayed immunoglobulin heterodimer, thereby producing a library of CDR-mutagenized phage-displayed immunoglobulin heterodimers.

2. The method of claim 1 wherein said 3' terminus has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122), or an oligonucleotide having a sequence complementary thereto.

3. The method of claim 1 wherein said 5' terminus has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123) or an oligonucleotide having a sequence complementary thereto.

4. The method of claim 1 wherein said template immunoglobulin heavy and light chain genes are obtained from a human.

5. The method of claim 1 wherein said CDR portion is CDR3.

6. The method of claim 1 wherein said formula is:

5'-GTGTATTATTGTGCGAGA$[NNS]_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 124).

7. The method of claim 1 wherein n is 16 in the formula $[NNS]_n$ (SEQ ID NO 120).

8. The method of claim 1 wherein said formula is:

5'-GTGTATTATTGTGCGAGA$[NNK]_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 125).

9. A method for producing a complementarity determining region (CDR)-mutagenized phage-displayed immunoglobulin heterodimer with altered antigen binding specificity, the method comprising the steps of:

1) amplifying a CDR portion of a template immunoglobulin variable domain gene selected from the group consisting of a template immunoglobulin heavy chain variable domain gene and a template immunoglobulin light chain variable domain gene, wherein said template immunoglobulin heavy and light chain genes have a framework region and said CDR portion and encode respective heavy and light chain variable domain polypeptides having a preselected antigen binding specificity to a first antigen, and wherein said amplifying is by polymerase chain reaction (PCR) using a PCR primer oligonucleotide for mutagenizing a preselected nucleotide region in said CDR portion to alter said preselected antigen binding specificity, thereby forming a library of amplified CDR-mutagenized immunoglobulin gene fragments, said PCR primer oligonucleotide having 3' and 5' termini and comprising:

a) a nucleotide sequence at the 3' terminus capable of hybridizing to a first framework region of said selected template immunoglobulin variable domain gene;

b) a nucleotide sequence at the 5' terminus capable of hybridizing to a second framework region of said selected template immunoglobulin variable domain gene; and

c) a nucleotide sequence between the 3' and 5' termini according to the formula selected from the group consisting of:

$[NNS]_n$ and $[NNK]_n$,

wherein N is independently any nucleotide, S is G or C and K is G or T, and n is 3 to about 24, the 3' and 5' terminal nucleotide sequences having a length of about 6 to 50 nucleotides, or an oligonucleotide having a sequence complementary thereto;

2) inserting individual members of the library of amplified CDR-mutagenized immunoglobulin gene fragments formed in step (1) into a dicistronic phagemid expression vector comprising immunoglobulin heavy and light chain variable domain genes that lack the immunoglobulin gene portion corresponding to the fragment to be inserted, wherein upon insertion said vector is capable of expressing heavy and light chain variable domain polypeptides encoded by said vector, thereby forming a library of dicistronic expression vectors containing amplified CDR-mutagenized immunoglobulin gene fragments;

3) expressing said immunoglobulin heavy and light chain genes in the library of dicistronic expression vectors formed in step (2) whereby said encoded heavy and light chain variable domain polypeptides assemble on the surface of a phage to form a phage-displayed immunoglobulin heterodimer, thereby producing a library of CDR-mutagenized phage-displayed immunoglobulin heterodimers; and

4) immunoreacting members of the library of CDR-mutagenized phage-displayed immunoglobulin heterodimers formed in step (3) on a preselected second antigen, said second antigen being different than said first antigen to allow for selection of a CDR-mutagenized phage-displayed immunoglobulin heterodimer with altered antigen binding specificity.

10. The method of claim **9** wherein said 3' terminus has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122), or an oligonucleotide having a sequence complementary thereto.

11. The method of claim **9** wherein said 5' terminus has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123) or an oligonucleotide having a sequence complementary thereto.

12. The method of claim **9** wherein said template immunoglobulin heavy and light chain genes are obtained from a human.

13. The method of claim **9** wherein said CDR portion is CDR3.

14. The method of claim **9** wherein said formula is

5'-GTGTATTATTGTGCGAGA[NNS]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 124).

15. The method of claim **9** wherein n is 16in the formula [NNS]$_n$ (SEQ ID NO 120).

16. The method of claim **9** wherein said formula is:

5'-GTGTATTATTGTGCGAGA[NNK]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 125).

17. A method for producing a soluble complementarity determining region (CDR)-mutagenized immunoglobulin heterodimer with altered immunoreactivity to a preselected antigen, the method comprising the steps of:

1) amplifying a CDR portion of a template immunoglobulin variable domain gene selected from the group consisting of a template immunoglobulin heavy chain variable domain gene and a template immunoglobulin light chain variable domain gene, wherein said template immunoglobulin heavy and light chain genes have a framework region and said CDR portion and encode respective heavy and light chain variable domain polypeptides immunoreactive with a preselected antigen, and wherein said amplifying is by polymerase chain reaction (PCR) using a PCR primer oligonucleotide for mutagenizing a preselected nucleotide region in said CDR portion to alter said immunoreactivity of said immunoglobulin heterodimer to the preselected antigen, thereby forming a library of amplified CDR-mutagenized immunoglobulin gene fragments, said PCR primer oligonucleotide having 3' and 5' termini and comprising:

a) a nucleotide sequence at the 3' terminus capable of hybridizing to a first framework region of said selected template immunoglobulin variable domain gene;

b) a nucleotide sequence at the 5' terminus capable of hybridizing to a second framework region of said selected template immunoglobulin variable domain gene; and

c) a nucleotide sequence between the 3' and 5' termini according to the formula selected from the group consisting of:

[NNS]$_n$ and [NNK]$_n$

wherein N is independently any nucleotide, S is G or C and K is G or T, and n is 3 to about 24, the 3' and 5' terminal nucleotide sequences having a length of about 6 to 50 nucleotides, or an oligonucleotide having a sequence complementary thereto;

2) inserting individual members of the library of amplified CDR-mutagenized immunoglobulin gene fragments formed in step (1) into a dicistronic phagemid expression vector comprising immunoglobulin heavy and light chain variable domain genes that lack the immunoglobulin gene portion corresponding to the fragment to be inserted, wherein upon insertion said vector is capable of expressing heavy and light chain variable domain polypeptides encoded by said vector, thereby forming a library of dicistronic expression vectors containing amplified CDR-mutagenized immunoglobulin gene fragments;

3) expressing said immunoglobulin heavy and light chain genes in the library of dicistronic expression vectors formed in step (2) whereby said encoded heavy and light chain variable domain polypeptides assemble on the surface of a phage to form a phage-displayed immunoglobulin heterodimer, thereby producing a library of CDR-mutagenized phage-displayed immunoglobulin heterodimers;

4) immunoreacting members of the library of CDR-mutagenized phage-displayed immunoglobulin heterodimers produced in step (5) on said preselected antigen to isolate an immunoglobulin heterodimer having altered immunoreactivity;

5) isolating said immunoreacted CDR-mutagenized phage-displayed immunoglobulin heterodimer obtained in step (4);

6) producing a soluble form of said immunoreacted form of immunoreacted CDR-mutagenized phage-displayed immunoglobulin heterodimer isolated in step (5); and

7) assaying said soluble form of immunoreacted CDR-mutagenized immunoglobulin heterodimer prepared in step (6) to identify a CDR-mutagenized immunoglobulin heterodimer with altered immunoreactivity to the preselected antigen.

18. The method of claim **17** wherein said assaying is determined by an increase in affinity to the preselected antigen.

19. The method of claim 18 wherein the affinity is greater than 10^5M^{-1} dissociation constant (K_d).

20. The method of claim 17 wherein said 3' terminus has the nucleotide sequence 5'-TGGGGCCAAGGGACCACG-3' (SEQ ID NO 122). or an oligonucleotide having a sequence complementary thereto.

21. The method of claim 17 wherein said 5' terminus has the nucleotide sequence 5'-GTGTATTATTGTGCGAGA-3' (SEQ ID NO 123) or an oligonucleotide having a sequence complementary thereto.

22. The method of claim 17 wherein said template immunoglobulin heavy and light chain genes are obtained from a human.

23. The method of claim 17 wherein said CDR portion is CDR3.

24. The method of claim 17 wherein said formula is

5'-GTGTATTATTGTGCGAGA[NNS]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 124).

25. The method of claim 17 wherein n is 16in the formula $[NNS]_n$ (SEQ ID NO 120).

26. The method of claim 17 wherein said formula is:

5'-GTGTATTATTGTGCGAGA[NNK]$_n$TGGGGCCAAGGGAC-CACG-3' (SEQ ID NO 125).

* * * * *

In vitro evolution of a T cell receptor with high affinity for peptide/MHC

Phillip D. Holler*, Philmore O. Holman*, Eric V. Shusta†, Sean O'Herrin‡, K. Dane Wittrup†§, and David M. Kranz*¶

Departments of *Biochemistry and †Chemical Engineering, University of Illinois, Urbana, IL 61801; and ‡Ben May Institute for Cancer Research, University of Chicago, Chicago, IL 60637

Communicated by Herman N. Eisen, Massachusetts Institute of Technology, Cambridge, MA, February 23, 2000 (received for review December 15, 1999)

T cell receptors (TCRs) exhibit genetic and structural diversity similar to antibodies, but they have binding affinities that are several orders of magnitude lower. It has been suggested that TCRs undergo selection *in vivo* to maintain lower affinities. Here, we show that there is not an inherent genetic or structural limitation on higher affinity. Higher-affinity TCR variants were generated in the absence of *in vivo* selective pressures by using yeast display and selection from a library of Vα CDR3 mutants. Selected mutants had greater than 100-fold higher affinity ($K_D \approx 9$ nM) for the peptide/MHC ligand while retaining a high degree of peptide specificity. Among the high-affinity TCR mutants, a strong preference was found for CDR3α that contained Pro or Gly residues. Finally, unlike the wild-type TCR, a soluble monomeric form of a high-affinity TCR was capable of directly detecting peptide/MHC complexes on antigen-presenting cells. These findings prove that affinity maturation of TCRs is possible and suggest a strategy for engineering TCRs that can be used in targeting specific peptide/MHC complexes for diagnostic and therapeutic purposes.

T cells recognize a foreign peptide bound to the MHC product through the $\alpha\beta$ heterodimeric receptor. The T cell receptor TCR) repertoire has extensive diversity created by the same gene rearrangement mechanisms used in antibody heavy- and light-chain genes (1). Most of the diversity is generated at the junctions of V and J (or diversity, D) regions that encode the complementarity-determining region three (CDR3) of the α and β chains (2). However, TCRs do not undergo somatic point mutations as do antibodies, and perhaps not coincidentally, TCRs also do not undergo the same extent of affinity maturation as antibodies. TCRs appear to have affinities that range from 10^5 to 10^7 M^{-1} whereas antibodies have affinities that range from 10^5 to 10^{10} M^{-1} (3, 4).

Whereas the absence of somatic mutation in TCRs may be associated with lower affinities, it has also been argued that there is not a selective advantage for a TCR to have higher affinity (5–7). In fact, the serial-triggering (6) and kinetic proofreading (7) models of T cell activation both suggest that very slow off-rates (associated with higher affinity) would be detrimental to the signaling process. On the other hand, the fastest off-rates that have been measurable have been associated with altered pMHC that exhibits antagonist activity (8–11). Whereas the narrow range of natural TCR affinities has provided some evidence for the relationships between off-rates and agonist/antagonist activity, there are also examples that appear to be inconsistent with these hypotheses (12, 13).

There are other possible explanations for why the T cell system maintains relatively low TCR:pMHC affinities *in vivo*. Peptides bound within the MHC groove display limited accessible surface (14), which may in turn limit the amount of free energy that can be generated in the interaction. On the other hand, raising the affinity of a TCR by directing the free energy toward the MHC helices would presumably lead to thymic deletion during negative selection (15). Even if such higher-affinity TCR could escape thymic deletion, they would likely not maintain the peptide specificity required for T cell responses.

It has not been possible to directly test these possibilities because the generation of TCRs with affinities above 10^7 M^{-1} has not been accomplished. In addition to allowing a kinetic basis of T cell triggering, high-affinity TCRs could be used to more easily explore the role of peptide in pMHC recognition, and as quantitative probes for the expression of pMHC on various target cells. Because *in vivo* selection schemes have not yielded TCRs with the intrinsic binding affinities of affinity-matured antibodies, in this report, we have used an *in vitro* method for the directed evolution of high-affinity TCRs. The method relies on the expression of a library of mutant single-chain (Vβ-linker-Vα) TCRs on the surface of yeast, as a fusion to the surface protein Aga-2 (16, 17). Our previous studies have shown that the yeast display system could be used to engineer variants of the 2C single-chain TCR (scTCR) that were more thermally stable and secreted at higher levels (17, 18). The stability mutants were isolated by subjecting the entire TCR gene to random mutagenesis and selecting for increased surface levels with anti-TCR antibodies (17). The mutations that increased stability resided at the Vα:Vβ interface or on the outside surface of Vβ in a region not involved in pMHC binding. To isolate TCR with higher affinity for pMHC, in the present study, we mutated only the CDR3α loop, which is at the center of the pMHC-binding site (19). Our efforts were guided by previous findings that this region contributed minimal binding free energy to the interaction of the 2C TCR with the pMHC ligand QL9/L^d (20), suggesting that productive interactions might be improved by focusing on this region. Remarkably, selection from a relatively small library (10^5 mutants) yielded many different TCRs with up to 100-fold increased affinity for QL9/L^d. The high-affinity TCRs retained a high degree of peptide specificity although there was some variation in fine specificity among the mutants. These findings suggest that the *in vitro* evolution process described here can be used to isolate TCRs with specificities that one defines by selection with appropriate pMHC ligands.

The high-affinity receptors in this study were derived by variation at the VJ junction, the same process that operates very effectively *in vivo* through gene rearrangements in T cells (2). The fact that we could readily isolate a diverse set of high-affinity TCR *in vitro* indicates that there is not a genetic or structural limitation to high-affinity receptors. This supports the view that inherently low affinities of TCRs found *in vivo* are caused by a lack of selection for higher affinity and perhaps a selection for lower affinity (5–7). Finally, the high-affinity TCR were used in monomeric form to detect pMHC on the surface of target cells,

Abbreviations: TCR, T cell receptor; scTCR, single-chain T cell receptor; CDR3, complementarity-determining regions three.

§Present address: Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

¶To whom reprint requests should be addressed. E-mail: d-kranz@uiuc.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.080078297. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.080078297

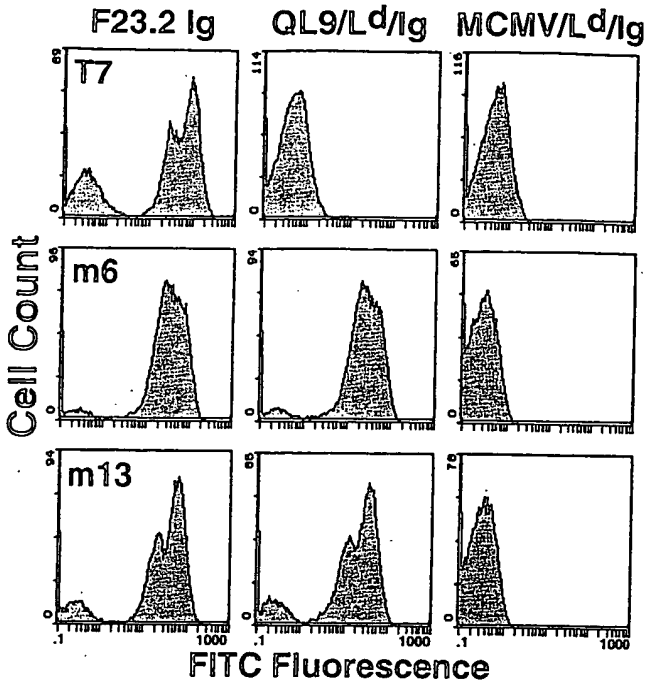
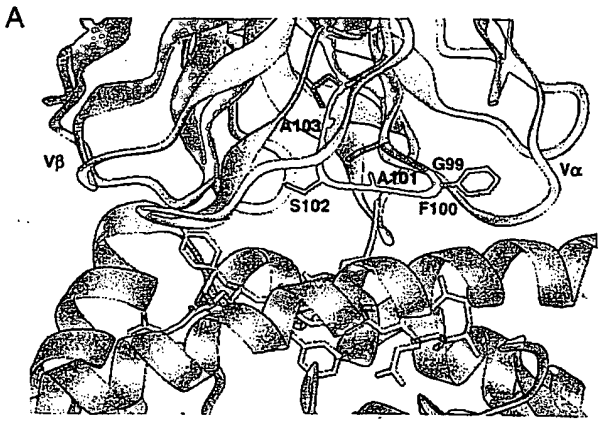


Fig. 1. Flow cytometric analysis of yeast cells that express wild-type and mutant 2C TCR on their surface. Yeast cells displaying wild-type (T7) and mutant (m6 and m13) scTCR were stained with anti-Vβ8 antibody F23.2 (120 nM), the specific alloantigenic peptide-MHC, QL9/L^d/Ig (40 nM), or a null peptide MCMV/L^d/Ig (40 nM). The peptides used in this study were QL9 (QLSPFPFDL), MCMV (YPHFMPTNL), and p2Ca (LSPFPFDL). Binding was detected by FITC-conjugated goat anti-mouse IgG F(ab')$_2$ and analyzed by flow cytometry. The negative population (e.g., seen with F23.2 staining) has been observed for all yeast-displayed proteins and is thought to be caused by cells at a stage of growth or induction that are incapable of expressing surface fusion protein (16, 17, 27).



B

Wild Type TCR		VαCDR3
	2C	^{93}SGFASAL^{104}

Mutant TCR	VαCDR3	Mutant TCR	VαCDR3
m1	SSYGNYL	m10	SLPPPLL
m2	SRRGHAL	m11	SIPTPSL
m3	SSRGTAL	m12	SNPPPLL
m4	SHPGTRL	m13	SDPPPLL
m5	SMFGTRL	m14	SSPPPRL
m6	SHQGRYL	m15	SAPPPIL
m7	SYLGLRL		
m8	SKHGIHL		
m9	SLTGRYL		

Fig. 2. Structure and sequences of the 2C TCR CDR3α. (*A*) X-ray crystallographic structure of the 2C/dEV8/K^b complex with CDR3α aa highlighted. Five residues of the 2C VαCDR3 that were randomized by PCR with a degenerate primer are shown in red. The adjacent CDR3 residues, Ser-93 and Leu-104 shown in blue, were retained in the yeast display library because they have been shown to be important in pMHC binding (17, 18, 20). (*B*) Alignment of aa sequences of mutant scTCRs isolated by yeast display and selection with QL9/L^d. Display plasmids were isolated from yeast clones after selection and sequenced to determine CDR3α sequences. Mutants m1, m2, m3, m4, m10, and m11 were isolated after the third round of sorting. All other mutants were isolated after the fourth round of sorting.

indicating that soluble forms of the TCR selected with the yeast display system can serve as probes for tumor-associated pMHC or other T cell-specific ligands.

Materials and Methods

Library Construction. The 2C single-chain TCR (scTCR) used as the scaffold for directed evolution (T7) contained six mutations (βG17E, βG42E, βL81S, αL43P, αW82R, and αI118N) that have been shown to increase the stability of the TCR but still allow pMHC binding (E.V.S., K.D.W., and D.M.K., unpublished results; and ref. 18). Mutagenic PCR of the T7 scTCR VαCDR3 was performed by using an AGA-2-specific upstream primer and a degenerate downstream primer 5'-CTTTTGTGCCGGATC-CAAATGTCAG(SNN)$_5$GCTCACAGCACAGAAGTACACG-GCCGAGTCGCTC-3'. Underlined bases indicate the positions of silent mutations introducing unique *Bam*HI and *Eag*I restriction sites. The purified PCR product was digested with *Nde*I and *Bam*HI and ligated to *Nde*I–*Bam*HI-digested T7/pCT302 (16–18). The ligation mixture was transformed into DH10B electrocompetent *Escherichia coli* (GIBCO/BRL), and transformants were pooled into 250-ml LB supplemented with ampicillin at 100 μg/ml and grown overnight at 37°C. Plasmid DNA was transformed into the yeast strain EBY100 by the method of Gietz and Schiestl (21).

Cell Sorting. The yeast library (22) was grown in 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids (Difco) at 30°C to an OD_{600} = 4.0. To induce surface scTCR expression, yeast were pelleted by centrifugation, resuspended to an OD_{600} = 1.0 in 2% galactose/0.67% yeast nitrogen base/1% Casamino acids, and incubated at 20°C for ≈24 h. In general, ≈10^7 cells per tube were incubated on ice for 1 h with 50 μl of QL9/L^d/IgG dimers (23) diluted in PBS (pH 7.4) supplemented with 0.5 mg/ml BSA. After incubation, cells were washed and labeled for 30 min with FITC-conjugated goat anti-mouse IgG F(ab')$_2$ (Kirkegaard & Perry) in PBS (pH 7.4) supplemented with 0.5 mg/ml BSA. Yeast were then washed and resuspended in PBS (pH 7.4) supplemented with 0.5 mg/ml BSA immediately before sorting. Cells exhibiting the highest fluorescence were isolated by using a Coulter 753 bench fluorescence-activated cell sorter. After isolation, sorted cells were expanded in 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids and induced in 2% galactose/0.67% yeast nitrogen base/1% Casamino acids for subsequent rounds of selection. A total of four sequential sorts were performed. The concentrations of QL9/L^d/IgG dimers used for staining were 50 μg/ml for sorts one to three and 0.5 μg/ml for the final sort. The percentages of total cells isolated from each sort were 5.55, 2.68, 2.56, and 0.58%, respectively. Aliquots of sorts three and four were plated on 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids to isolate individual clones which were analyzed by flow cytometry by using a Coulter Epics XL instrument.

Soluble scTCR Production. The T7 and m6 scTCR genes were excised from pCT302 *Nhe*I–*Xho*I and ligated into *Nhe*I–*Xho*I

digested pRSGALT, a yeast expression plasmid (18). Ligations were transformed into DH10B electrocompetent *E. coli* (GIBCO/BRL). Plasmid DNA was isolated from bacterial cultures and transformed into the *Saccharomyces cerevisiae* strain, BJ5464 *(α ura3–52 trp1 leu2Δ1 his3Δ200 pep4::HIS3 prb1Δ1.6R can1 GAL)* (18). Yeast clones were grown in 1 liter of 2% dextrose/0.67% yeast nitrogen base/1% Casamino acids/20 mg/liter Trp for 48 h at 30°C. To induce scTCR secretion, cells were pelleted by centrifugation at 4,000 × *g*, resuspended in 1 liter of 2% galactose/0.67% yeast nitrogen base/1% Casamino acids/20 mg/l Trp supplemented with 1 mg/ml BSA, and incubated for 72 h at 20°C. Culture supernatants were harvested by centrifugation at 4,000 × *g*, concentrated to ≈50 ml, and dialyzed against PBS (pH 8.0). The six His-tagged scTCRs were purified by native nickel affinity chromatography [Ni-NTA Superflow, Qiagen (Chatsworth, CA); 5 mM and 20 mM imidazole (pH 8.0) wash; 250 mM imidazole elution] (18).

Cell-Binding Assays. The binding of soluble scTCRs to QL9/L^d was monitored in a competition format as described (20, 24). Peptide-loaded T2-L^d cells (3×10^5 per well) were incubated for 1 h on ice in the presence of ^{125}I-labeled anti-L^d Fabs (30–5-7) and various concentrations of scTCRs. Bound and unbound [^{125}I] 30–5-7 Fabs were separated by centrifugation through olive oil/dibutyl phthalate. Inhibition curves were constructed to determine inhibitor concentrations yielding 50% maximal inhibition. Dissociation constants were calculated by using the formula of Cheng and Prusoff (25). To monitor direct binding of scTCRs to cell-bound pMHC, peptide-loaded T2-L^d cells (5×10^5 per tube) were incubated for 40 min on ice with biotinylated soluble scTCRs followed by staining for 30 min with streptavidin-phycoerythrin (PharMingen). Cellular fluorescence was detected by flow cytometry.

Results and Discussion

To examine if it is possible to generate higher-affinity TCR that would retain peptide specificity, we subjected a TCR to a process of directed *in vitro* evolution. Phage display (26) has not yet proven successful in the engineering of single-chain TCRs (scTCRs, Vβ-linker-Vα) despite the extensive structural similarity between antibody and TCR V regions. However, we recently showed that a scTCR could be displayed on the surface of yeast (17) in a system that has proven successful in antibody engineering (16, 27). A temperature-stabilized variant (T7) (18) of the scTCR from the cytotoxic T lymphocyte clone 2C was used in the present study. Cytotoxic T lymphocyte clone 2C recognizes the alloantigen L^d with a bound octamer peptide called p2Ca, derived from the enzyme 2-oxoglutarate dehydrogenase (28). The nonameric variant QL9 is also recognized by cytotoxic T lymphocyte 2C, but with 10-fold higher affinity by the 2C TCR (29). Alanine scanning mutagenesis showed that the CDR3α loop contributed minimal free energy to the binding interaction (20), even though structural studies have shown that CDR3α of the 2C TCR is near the peptide and it undergoes a conformational change to accommodate the pMHC complex (19). Thus, we focused our mutagenesis efforts on five residues that form the CDR3α loop.

A library of 10^5 independent TCR-CDR3α yeast mutants was subjected to selection by flow cytometry with a fluorescently labeled QL9/L^d ligand (23). After four rounds of sorting and growth, 15 different yeast colonies were examined for their ability to bind the ligand, in comparison to the scTCR variant T7, which bears the wild-type CDR3α sequence (Fig. 1 and data not shown). The anti-Vβ8.2 antibody F23.2 which recognizes residues in the CDR1 and CDR2 was used as a control to show that wild-type scTCR-T7 and scTCR mutants (m6 and m13, in Fig. 1, and others, data not shown) each had approximately equivalent surface levels of the scTCR (Fig. 1). In contrast, the soluble QL9/L^d ligand bound very well to each mutant yeast clone but not to wild-type scTCR-T7. The MCMV/L^d complex, which is not recognized by cytotoxic T lymphocyte clone 2C, did not bind

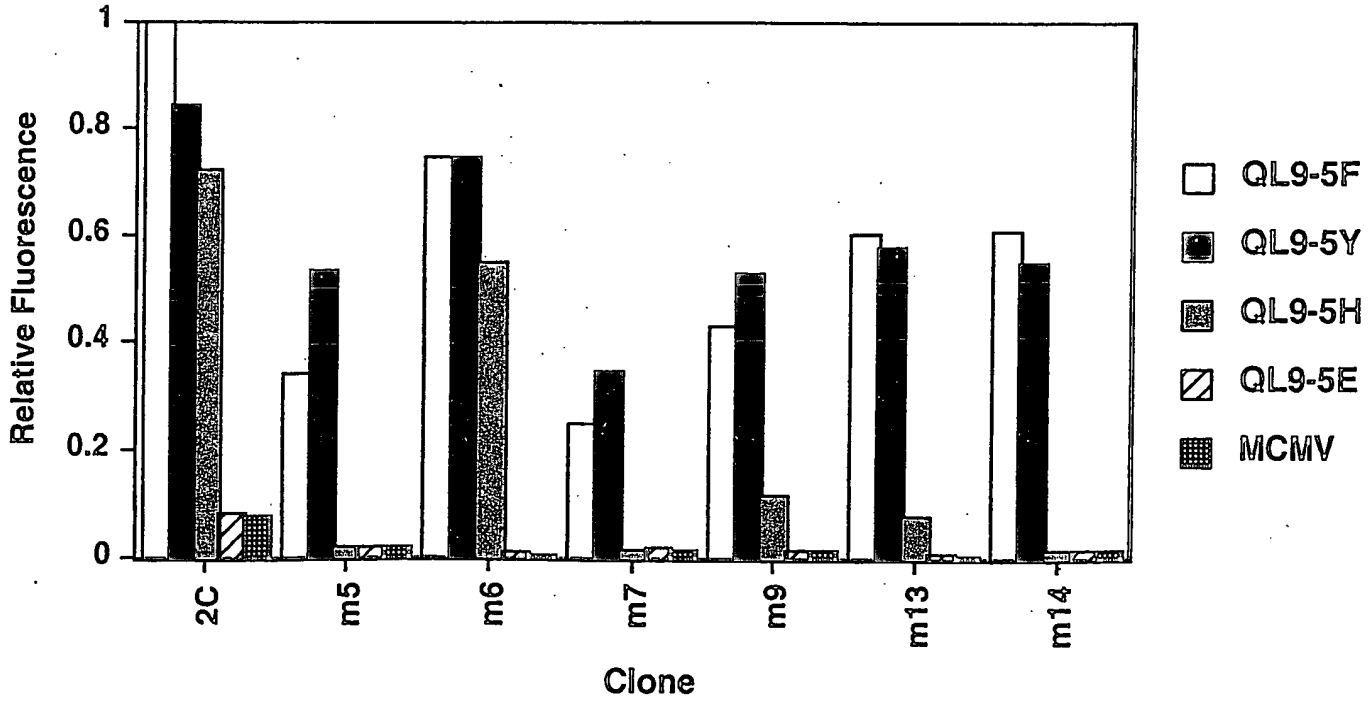


Fig. 3. Fine specificity analysis of mutant scTCR binding to different QL9 variant peptides bound to L^d. The original T cell clone 2C and various yeast clones were analyzed by flow cytometry for binding to L^d/Ig dimers loaded with wild-type QL9 (QL9–5F), position 5 variants of QL9 (QL9–5Y, QL9–5H, and QL9–5E) or MCMV. Binding was detected with FITC-labeled goat anti-mouse IgG. Relative fluorescence was measured by comparison with mean fluorescence values of 2C cells or yeast cells stained with anti-Vβ8 antibody F23.2.

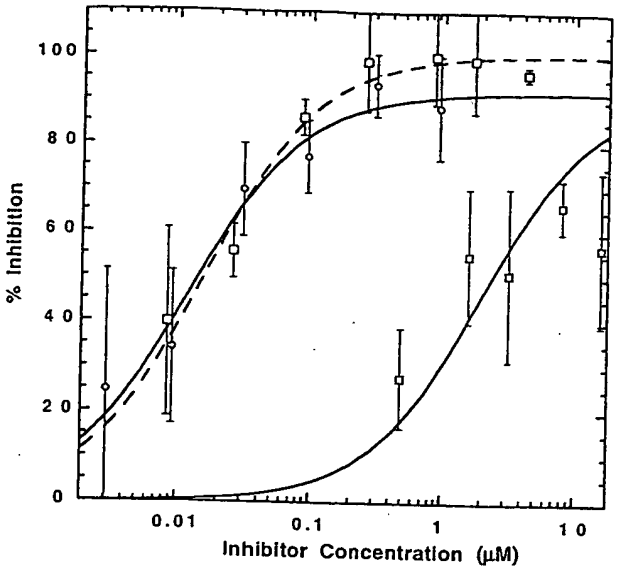
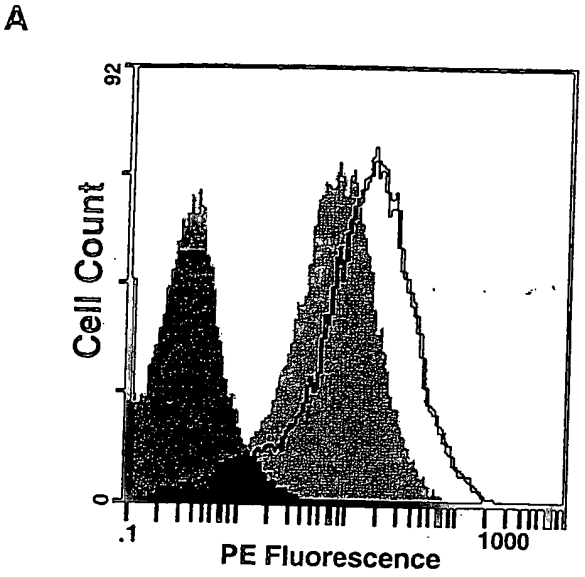


Fig. 4. QL9/L^d binding by soluble scTCRs. T2-L^d cells loaded with QL9 were incubated with ^{125}I-labeled anti-L^d Fab fragments (30–5-7) and various concentrations of unlabeled Fab (□), scTCR-T7 (⧈), or mutant scTCR-m6 (⊙). Bound and unbound [^{125}I]30–5-7 Fab fragments were separated by centrifugation through olive oil/dibutyl phthalate. Binding of ^{125}I-labeled anti-L^d Fab fragments to T2-L^d cells loaded with the control peptide MCMV was not inhibited even at the highest concentrations of scTCRs (data not shown).



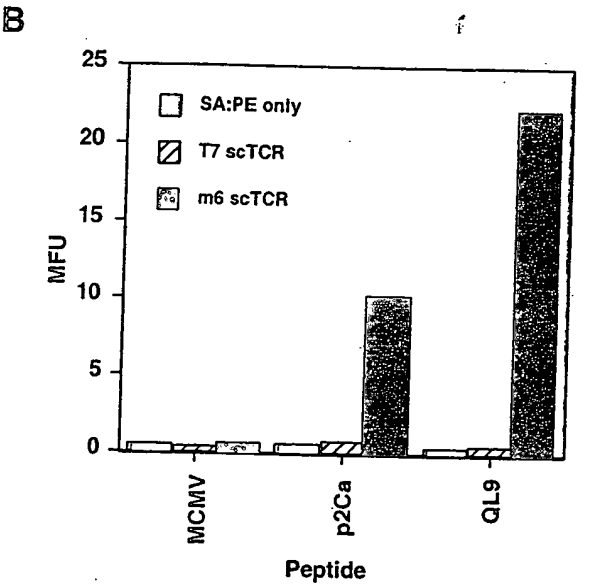


Fig. 5. Flow cytometric analysis of the binding of scTCR/biotin to cell surface peptide/MHC. Peptide-loaded T2-L^d cells were incubated with biotinylated m6 scTCR (~0.3 μM) or T7 scTCR (~1.6 μM) scTCR followed by streptavidin-PE and analyzed by flow cytometry. (*A*) Flow cytometry histograms of T2-L^d cells loaded with QL9 (unshaded), p2Ca (light shade), or MCMV (dark shade) and stained with m6 scTCR/biotin. (*B*) Mean fluorescent units (MFU) of T2-L^d cells loaded with QL9, p2Ca, or MCMV and stained with either secondary SA-PE only, T7 scTCR/biotin + SA-PE, or m6 scTCR/biotin + SA-PE.

to the scTCR mutants or wild-type scTCR-T7, indicating that the scTCR mutants retained peptide specificity.

The CDR3α sequences of the 15 mutants all differed from the 2C TCR (Fig. 2). It was readily apparent (and confirmed by a BLAST alignment algorithm) that the sequences could be aligned into two motifs. One motif contained Gly in the middle of the five residue stretch whereas the other motif contained three tandem Pro. Evidence that all three Pro are important in generating the highest affinity site is suggested by results with mutant m11. Mutant m11 contained only two of the three Pro and exhibited reduced binding compared with the triple-Pro mutants (data not shown). The Gly-containing mutants appeared to have preferences for positive-charged residues among the two residues to the carboxyl side (7/9) and aromatic and/or positive-charged residues among the two residues to the amino side (4/9 and 5/9). The selection for a glycine residue at position 102 in the motif may indicate that the CDR3α loop required conformational flexibility around this residue to achieve increased affinity. This is consistent with the large (6-Å) conformational difference observed between the CDR3α loops of the liganded and unliganded 2C TCR (19). It is also interesting that Gly is the most common residue at the V(D)J junctions of antibodies and the presence of a Gly has recently been associated with increased affinity in the response to the (4-hydroxy-3-nitrophenyl) acetyl hapten (30).

In contrast to the isolates that contain Gly, the selection for a Pro-rich sequence at the tip of the CDR3α loop may suggest that these TCR exhibit a more rigid conformation that confers higher affinity. The x-ray crystallographic structures of a germline antibody of low affinity compared with its affinity-matured derivative showed that the high-affinity state was associated with stabilization of the antibody in a configuration that accommodated the hapten (31). Similarly, the NMR solution structure of a scTCR that may be analogous to the germ-line antibody showed that the CDR3α and β loops both exhibited significant mobility (32). Recent thermodynamic studies of TCR:pMHC interactions have also suggested the importance of conformational changes in binding (33, 34). Structural and thermodynamic studies of the TCR mutants that we report here should allow us to examine if the two CDR3α motifs (Gly- versus Pro-rich) might differ in the mechanism by which they confer higher affinity.

Although the scTCR mutants did not bind the null peptide/L^d complex MCMV/L^d, it remained possible that the increase in affinity might be accompanied by a change in fine specificity. To examine this issue, we used QL9 position 5 (Phe) peptide variants that have been shown previously to exhibit significant differences in their binding affinity for the wild-type 2C TCR (35). The binding of these pMHC to various TCR mutants on the yeast surface and clone 2C were measured by flow cytometry. As

shown in Fig. 3, the native TCR on 2C is capable of binding QL9 variants that contain either Tyr or His at position five but not Glu. Each of the higher-affinity TCR mutants retained their ability to recognize the conserved Tyr-substituted peptide and they were likewise incapable of recognizing the Glu-substituted peptide. However, several of the TCR mutants (m6, m9, and m13) bound to the His-substituted peptide (albeit to different extents) whereas other mutants (m5, m7, and m14) did not bind the peptide. Thus, the CDR3α loop can influence the peptide fine specificity of recognition but it is not the only region of the TCR involved. The effect on peptide specificity could be through direct interaction of CDR3α residues with the variant peptide, as suggested from earlier studies involving CDR3-directed selections (36, 37). Alternatively, binding free energy may be directed at peptide-induced changes in the L^d molecule itself. The latter possibility is perhaps more likely in the case of the 2C TCR:QL9/L^d interaction, because position five of QL9 has been predicted to point toward the L^d groove (35, 38). The fine-specificity analysis also shows that it is possible to engineer TCRs with increased, or at least altered, specificity for cognate peptides. Thus, directed evolution of only a short region (CDR3α) of a single TCR allows the design of TCR variants with altered peptide-binding specificities.

To determine the magnitude of the affinity increases associated with a selected CDR3α mutant, the wild-type T7 scTCR and the m6 scTCR were expressed as soluble forms in a yeast secretion system. Purified scTCR preparations were compared for their ability to block the binding of a ^{125}I-labeled anti-L^d Fab fragments to QL9 or MCMV loaded onto L^d on the surface of T2-L^d cells. As expected, neither T7 nor m6 scTCR were capable of inhibiting the binding of ^{125}I-Fab fragments to T2-L^d cells up-regulated with the MCMV peptide (data not shown). However, both T7 and m6 were capable of inhibiting the binding of anti-L^d Fab fragments to QL9/L^d (Fig. 4). The m6 scTCR variant was as effective as unlabeled Fab fragments in inhibiting binding, whereas the T7 scTCR was 160-fold less effective (average of 140-fold difference among four independent titrations). The K_D values of the scTCR for QL9/L^d were calculated from the inhibition curves to be 1.5 μM for T7 and 9.0 nM for m6. The value for T7 is in close agreement with the 3.2 μM K_D reported for the 2C scTCR (39). These findings show that the yeast system, combined with CDR3α-directed mutagenesis, is capable of selecting mutants with 100-fold higher intrinsic binding affinities for a pMHC ligand.

If the soluble scTCR has a high affinity for its pMHC ligand, then it may be useful, like antibodies, as a specific probe for cell surface-bound antigen. To test this possibility, the soluble T7 and m6 scTCR were biotinylated and the labeled scTCR were incubated with T2-L^d cells loaded with QL9, p2Ca, or MCMV. The m6 scTCR, but not the T7 scTCR, yielded easily detectable staining of the T2 cells that had been incubated with QL9 or p2Ca (Fig. 5 *A* and *B*). It is significant that p2Ca-up-regulated cells were also readily detected by m6 scTCR, because p2Ca is the naturally processed form of the peptide recognized by the alloreactive clone 2C and it has an even lower affinity than the QL9/L^d complex for the 2C TCR (29). However, it remains to be determined if the levels of pMHC derived from endogenous antigen processing are sufficient to allow detection by using soluble TCR as probes. It is reasonable to predict that, in some cases, the level will be too low to distinguish from background by using standard flow cytometry procedures.

The high-affinity receptors described in our study were derived by variation at the VJ junction, the same process that operates very effectively *in vivo* through gene rearrangements in T cells (2). The fact that we could readily isolate a diverse set of high-affinity TCR *in vitro* indicates that there is not a genetic or structural limitation to high-affinity receptors. This supports the view that inherently low affinities of TCRs found *in vivo* are caused by a lack of selection for higher affinity and perhaps a selection for lower affinity (5–7). In this respect, the higher-affinity TCRs now provide the reagents for directly testing hypotheses about the effects of affinity on T cell responses (4–7). It is interesting to note that similar arguments have been used to suggest that the kinetic properties of antibodies may also set an *in vivo* "affinity ceiling," above which there may not be a selective advantage to B cells (40).

In addition to their utility for testing T cell responses, high-affinity TCRs can be engineered like antibodies to yield high-affinity, antigen-specific probes. Soluble versions of the high-affinity receptor can directly detect specific peptide/MHC complexes on cells (Fig. 5). Thus, these engineered proteins have potential, for example, as tumor cell diagnostics, or on conjugation with cytotoxins, potential agents for cancer therapy.

We thank Michele Kieke for providing 2C scTCR genes cloned in the yeast display vector and Gary Durack of the University of Illinois Biotechnology Center Flow Cytometry Facility for assistance and advice. We also thank Jeff Bluestone and Ian Wilson for comments and discussion on the manuscript and results. This work was supported by National Institutes of Health Grant RO1 GM55767.

1. Tonegawa, S. (1988) *Biosci. Rep.* 8, 3–26.
2. Davis, M. M. & Bjorkman, P. J. (1988) *Nature (London)* 334, 395–402.
3. Eisen, H. N., Sykulev, Y. & Tsomides, T. J. (1996) *Adv. Protein Chem.* 49, 1–56.
4. Davis, M. M., Boniface, J. J., Reich, Z., Lyons, D., Hampl, J., Arden, B. & Chien, Y. (1998) *Annu. Rev. Immunol.* 16, 523–544.
5. Sykulev, Y., Cohen, R. J. & Eisen, H. N. (1995) *Proc. Natl. Acad. Sci. USA* 92, 11990–11992.
6. Valitutti, S., Muller, S., Cella, M., Padovan, E. & Lanzavecchia, A. (1995) *Nature (London)* 375, 148–151.
7. Rabinowitz, J. D., Beeson, C., Lyons, D. S., Davis, M. M. & McConnell, H. M. (1996) *Proc. Natl. Acad. Sci. USA* 93, 1401–1405.
8. Rabinowitz, J. D., Beeson, C., Wulfing, C., Tate, K., Allen, P. M., Davis, M. M. & McConnnell, H. M. (1996) *Immunity* 5, 125–135.
9. Sloan-Lancaster, J. & Allen, P. M. (1996) *Annu. Rev. Immunol.* 14, 1–27.
10. Alam, S. M., Travers, P. J., Wung, J. L., Nasholds, W., Redpath, S., Jameson, S. C. & Gascoigne, R. J. (1996) *Nature (London)* 381, 616–620.
11. Kersh, G. J., Kersh, E. N., Fremont, D. H. & Allen, P. M. (1998) *Immunity* 9, 817–826.
12. Al-Ramadi, B. K., Jelonek, M. T., Boyd, L. F., Margulies, D. H. & Bothwell, A. L. M. (1995) *J. Immunol.* 155, 662–673.
13. Sykulev, Y., Vugmeyster, Y., Brunmark, A., Ploegh, H. L. & Eisen, H. N. (1998) *Immunity* 9, 475–483.
14. Bjorkman, P. J. (1997) *Cell* 89, 167–170.
15. Bevan, M. J. (1997) *Immunity* 7, 175–178.
16. Boder, E. T. & Wittrup, K. D. (1997) *Nat. Biotechnol.* 15, 553–557.
17. Kieke, M. C., Shusta, E. V., Boder, E. T., Teyton, L., Wittrup, K. D. & Kranz, D. M. (1999) *Proc. Natl. Acad. Sci. USA* 96, 5651–5656.
18. Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M. & Wittrup, K. D. (1999) *J. Mol. Biol.* 292, 949–956.
19. Garcia, K. C., Degano, M., Pease, L. R., Huang, M., Peterson, P., Teyton, L. & Wilson, I. A. (1998) *Science* 279, 1166–1172.
20. Manning, T. C., Schlueter, C. J., Brodnicki, T. C., Parke, E. A., Speir, J. A., Garcia, K. C., Teyton, L., Wilson, I. A. & Kranz, D. M. (1998) *Immunity* 8, 413–425.
21. Gietz, R. D., Schiestl, R. H., Willems, A. & Woods, R. A. (1995) *Yeast* 11, 355–360.
22. Shusta, E. V., VanAntwerp, J. & Wittrup, K. D. (1999) *Curr. Opin. Biotechnol.* 10, 117–122.
23. Dal Porto, J., Johansen, T. E., Catipovic, B., Parfiit, D. J., Tuveson, D., Gether, U., Kozlowski, S., Fearon, D. T. & Schneck, J. P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 6671–6675.
24. Sykulev, Y., Brunmark, A., Jackson, M., Cohen, R. J., Peterson, P. A. & Eisen, H. N. (1994) *Immunity* 1, 15–22.
25. Cheng, Y. C. & Prusoff, W. H. (1973) *Biochem. Pharmacol.* 22, 3099–3108.
26. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. & Winter, G. (1991) *Nature (London)* 352, 624–628.
27. Kieke, M. C., Cho, B. K., Boder, E. T., Kranz, D. M. & Wittrup, K. D. (1997) *Protein Eng.* 10, 1303–1310.
28. Udaka, K., Tsomides, T. J., Walden, P. P., Fukusen, N. & Eisen, H. N. (1993) *Proc. Natl. Acad. Sci. USA* 90, 11272–11276.

29. Sykulev, Y., Brunmark, A., Tsomides, T. J., Kageyama, S., Jackson, M., Peterson, P. A. & Eisen, H. N. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 11487–11491.
30. Furukawa, K., Akasako-Furukawa, A., Shirai, H., Nakamura, H. & Azuma, T. (1999) *Immunity* **11**, 329–338.
31. Wedemayer, G. J., Patten, P. A., Wang, L. H., Schultz, P. G. & Stevens, R. C. (1997) *Science* **276**, 1665–1669.
32. Hare, B. J., Wyss, D. F., Osburne, M. S., Kern, P. S., Reinherz, E. L. & Wagner, G. (1999) *Nat. Struct. Biol.* **6**, 574–581.
33. Willcox, B. E., Gao, G. F., Wyer, J. R., Ladbury, J. E., Bell, J. I., Jakobsen, B. K. & Anton van der Merwe, P. (1999) *Immunity* **10**, 357–365.
34. Boniface, J. J., Reich, Z., Lyons, D. S. & Davis, M. M. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 11446–11451.
35. Schlueter, C. J., Manning, T. C., Schodin, B. A. & Kranz, D. M. (1996) *J. Immunol.* **157**, 4478–4485.
36. Sant'Angelo, D. B., Waterbury, G., Preston-Hurburt, P., Yoon, S. T., Medzhitov, R., Hong, S. & Janeway, C. A. (1996) *Immunity* **4**, 367–376.
37. Jorgensen, J. L., Esser, U., de St. Groth, B. F., Reay, P. A. & Davis, M. M. (1992) *Nature (London)* **355**, 224–230.
38. Speir, J. A., Garcia, K. C., Brunmark, A., Degano, M., Peterson, P. A., Teyton, L. & Wilson, I. A. (1998) *Immunity* **8**, 553–562.
39. Manning, T. C., Parke, E. A., Teyton, L. & Kranz, D. M. (1999) *J. Exp. Med.* **189**, 461–470.
40. Foote, J. & Eisen, H. N. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 1254–1256.